A Publication of the American Psychological Association

Journal of Abnormal and Social Psychology

Founded by Morton Prince, 1906

CONTENTS

Articles

Concept Attainment of Schizophrenics on Materials Involving Social Approval and Disapproval 355
David Moriarty and Solis L. Kates

The Influence of Positive Interpersonal Relations on Adjustment and Effectiveness in Rifle Teams 365
Joseph E. McGrath

Individual Differences, Situational Variables, and Personality Research 376
Irwin G. Sarason

Effects of Ambiguity of Environmental Cues upon Learning to Imitate 381
John W. McDavid

The Performance of Process and Reactive Schizophrenics and Brain Injured Subjects on a Conceptual Task 387
Thomas A. Tutko and Janet Taylor Spence

The Effectiveness of "Overheard" Persuasive Communications 395
Elaine Walster and Leon Festinger

An Experimental Study of Psychological Resistances to Fear Arousing Communications 403
Irving L. Janis and Robert P. Terwilliger

Physiological and Subjective Responses to Pain-Producing Stimulation under Hypnotically-Suggested and Waking-Imagined "Analgesia" 411
Theodore Xenophon Barber and Earl W. Hahn, Jr.

Case Report

Understanding Language without Ability to Speak: A Case Report 419
Eric H. Lenneberg

Critique and Notes

Two Anxiety Factors in Psychiatric Patients 426
Arnold H. Buss

Further Validation of the Alpert-Haber Achievement Anxiety Test 427
William N. Dember, Frank Nahrn, and Francis Joseph Miller

The Effectiveness of Social Reinforcement Following Two Conditions of Social Deprivation 429
Harold W. Stevenson and Richard D. Odom

A Cross-Validation of Twelve MMPI Indices of Hostility and Control 431
Edwin I. Megargee and Gerald A. Mendelsohn

Acceptance by Others and Its Relationship to Acceptance of Self and Others: A Repeat of Fey's Study 436
John E. Williams

This is the last issue of Volume 65.
Title page and Contents for the Volume appear herein.

Volume 65 DECEMBER 1962 Number 6

Executive Editor
JOHN E. BOREN

Editorial Board
D. G. ARONS
J. E. ARNASON
N. H. AZRIN
L. W. BIJOU
C. S. BLOUGH
J. V. BRADY
W. W. CUMMING
F. B. DAVIS
J. A. DINSMOOR
G. S. FERSTER
J. D. FOWLEY
L. R. GOLLUB
J. M. HAMLISTON
R. F. HENDERSON
R. J. HERRNSTEIN
J. G. HOLLAND
R. T. KELLEHER
F. S. KELLER
V. G. LAURER
O. R. LINDSLEY
C. B. LOVE
W. M. MORSE
W. N. SCHOENFELD
M. SIDMAN
B. F. SKINNER
T. VERHAVE
B. WEISS

Journal of the Experimental Analysis of Behavior

Some recent articles of interest to readers of the *JAEP*:

STRATE, A. W., STRATE, CORNELIA K., SCHOTT, R. L., and WOLF, MONTROSE. The conditioning of textual responses during "extinction" reinforcers.

WEISER, HAROLD. Some effects of response cost upon human operant behavior.

ANTHON, [illegible] and HAUGHTON, E. Control of the behavior of schizophrenic patients by food.

GOLDIAMOND, ISRAEL. Machine definition of ongoing silent and oral reading rate.

KEPPEL, GEOFFREY, and SHAFFER, DAVID. The operant conditioning of con-creation.

BOREN, DONALD M. Laboratory control of thumbsucking by withdrawal and re-presentation of reinforcement.

LONG, EUGENE R. Additional techniques for producing multiple-schedule control in children.

SKINNER, B. F. Two "synthetic social relations."

A personal subscription for 1963 (Vol. 6) is available to members of recognized professional societies at the special rate of $8.50. to students at the rate of $5.50. If you do not care to subscribe yourself, you may wish to call this journal to the attention of your librarian. Regular rates are $15.00 per year, and the first five volumes are still available at this rate if your librarian wishes a complete file.

All orders and business correspondence should be addressed to EAV DEPARTMENT, *Department of Psychology*, Lindley Hall, Indiana University, Bloomington, Indiana.

Volume 65 December 1962 Number 6

Journal of Abnormal and Social Psychology

Editor: DANIEL KATZ, *University of Michigan*

Associate Editor: ROBERT B. ZAJONC, *University of Michigan*

CONSULTING EDITORS

JOSEPH B. ADELSON
University of Michigan

GORDON W. ALLPORT
Harvard University

NORMAN H. ANDERSON
University of California, Los Angeles

LEONARD BERKOWITZ
University of Wisconsin

JACK W. BREHM
Duke University

URIE BRONFENBRENNER
Cornell University

DONALD R. BROWN
Bryn Mawr College

LAUNOR F. CARTER
System Development Corporation

DWIGHT W. CHAPMAN, JR.
Vassar College

MORTON DEUTSCH
Bell Telephone Laboratories

SEYMOUR FESHBACH
University of Pennsylvania

DONALD W. FISKE
University of Chicago

NATHANIEL L. GAGE
University of Illinois

NICHOLAS HOBBS
Peabody College for Teachers

MURRAY HORWITZ
New York University

IRVING L. JANIS
Yale University

HAROLD H. KELLEY
University of California, Los Angeles

HERBERT C. KELMAN
Harvard University

DANIEL J. LEVINSON
Harvard Medical School

DAVID C. MCCLELLAND
Harvard University

WILLIAM J. MCGUIRE
Columbia University

ROBERT B. MACLEOD
Cornell University

HAROLD L. RAUSH
National Institute of Mental Health

NEVITT SANFORD
Stanford University

IRVING SARNOFF
New York University

WILLIAM A. SCOTT
University of Colorado

JANET TAYLOR SPENCE
Veterans Administration Hospital, Iowa City

JOSEPH VEROFF
University of Michigan

JUSTIN L. WEISS
Massachusetts Mental Health Center

PUBLISHED BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, Inc.

The JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY is devoted to both abnormal and social psychology. Its emphasis is on basic research and theory rather than the techniques and arts of practice. Abnormal psychology is broadly defined to include papers contributing to fundamental knowledge of the pathology, dynamics, and development of personality or individual behavior, including deterioration with age and disease. Articles concerned with psychodiagnostic techniques are evaluated with respect to their contribution to an understanding of the psychological principles of diagnosis, those concerned with psychotherapy are judged in terms of their contribution to an understanding of the therapeutic process. Case reports are considered in terms of their heuristic value. Reports which serve to pose or to clarify important theoretical problems, or which promise to be useful in teaching, are often published. From the social area this JOURNAL gives preference to papers contributing to basic knowledge of interpersonal relations, and of group influences on the pathology, dynamics, and development of individual behavior.

The JOURNAL is issued monthly, two volumes per year. The subscription price per volume is $10.00, or $20.00 annually (foreign, $20.50). Single issues are $2.00.

Original manuscripts should be sent to the Editor:

DANIEL KATZ
University of Michigan
Box 1248
Ann Arbor, Michigan

Manuscripts must be accompanied by an abstract of 100–120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts appeared in the *American Psychologist* (1961, 16, 833), or may be obtained from the Editor or from the APA Central Office.

HELEN ORR	ELIZABETH S. REED	JUDY SMITH
Managing Editor	Advertising Manager	Editorial Assistant

Subscriptions, address changes, and business communications should be sent to:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1333 Sixteenth Street, N.W.
Washington 6, D. C.

Changes of address must reach the Subscription Office by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

PUBLISHED BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
Prince and Lemon Sts., Lancaster, Pa.
and 1333 Sixteenth Street N.W., Washington 6, D. C.

Second-class postage paid at Lancaster, Pa., and at additional mailing offices

Copyright © 1962 by the American Psychological Association, Inc.

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 355-364

CONCEPT ATTAINMENT OF SCHIZOPHRENICS ON MATERIALS INVOLVING SOCIAL APPROVAL AND DISAPPROVAL[1]

DAVID MORIARTY AND SOLIS L. KATES

University of Massachusetts

The purpose of this study is to investigate concept attainment on social materials in poor premorbid schizophrenics, good premorbid schizophrenics, and normals. The conceptual tasks involve the use of pictorial representations of interpersonal scenes suggesting approval and matched scenes suggesting disapproval.

The theories of Sullivan (1956), Cameron (1947), and Goldstein (in his more recent formulation of 1959) agree that schizophrenia is a reaction to severe interpersonal disturbance in infancy and early childhood, with a resulting sensitivity to interpersonal stimuli, particularly to stimuli that threaten personal security. Their writings suggest first that schizophrenics will perform less efficiently than normals on problem materials involving social content. Their writings imply, second, that schizophrenics will show greater impairment than normals on problem materials involving interpersonal censure than on more neutrally toned materials.

Included among the relevant studies apparently confirming the first inference above are those of Davis and Harrington (1957) and of Whiteman (1954). Davis and Harrington found that a schizophrenic group does as well as a normal group on nonhuman content tasks, but shows impaired behavior on a task involving human content. Whiteman's study indicated that normal subjects are superior to schizophrenics on both formal and social conceptualizations. Further, when the schizophrenic and normal subjects are matched in performance on the formal concept tasks, there is a significant decrement manifested by the schizophrenics on the social concept tasks. These studies may be interpreted as supporting the inference of selective decrements in schizophrenic conceptualization behavior on social tasks. They afford some evidence in favor of a limited conceptual inability in schizophrenics. Yet there is need of further replication.

Pertinent to the second inference, that schizophrenics show selective deficits on threatening socially toned material, is a study by Dunn (1954). He found that schizophrenics are significantly less effective than normals in discriminating between the modifications of a scolding scene, indicating that schizophrenics are especially liable to interpersonal threat in the form of censure. Heath (1956) showed that schizophrenics perform less adequately on dissected sentences concerned with threatening interpersonal themes than on dissected sentences dealing with nonthreatening interpersonal themes. Regrettably, a normal control group was not used. A study by Webb (1955) indicated that verbally censured schizophrenic subjects show conceptual deficit when compared with control schizophrenic subjects. Again, there were no normal subjects included. If other studies, which are necessary, indicate a selective conceptual impairment in schizophrenics when dealing with socially threatening materials—not also manifested by normals matched on formal conceptual tasks—then the notion of a limited conceptual inability in schizophrenics rather than a generalized one gains further support.

Recently, the premorbid history of the schizophrenic patients has been used for dividing those patients into two more or less homogeneous groups for experimental purposes (Phillips, 1953). In Dunn's (1954) study, only the poor premorbid schizophrenics are found to manifest a significant decrement in visual discrimination while the good premorbid schizophrenics do not differ from the normals. Bleke's (1955) investigation showed that censure leads to significantly greater in-

[1] Based in part upon a master's thesis completed by the senior author under the direction of the junior author.

terference effects in recall in poor premorbid schizophrenics than in good premorbid schizophrenics and in normals. The Phillips scale evaluates quantitatively the adequacy of the patient's social and sexual premorbid history, and can be considered essentially as a measure of his withdrawal from social life or his social disarticulation. It would follow logically that the poor premorbid schizophrenics, the patients with the greatest social disarticulation, would show the greatest impairment in conceptual ability upon tasks involving both social approval and disapproval themes. It would also follow that the poor premorbids would manifest the greatest decrement in conceptual ability in the transition from social approval tasks to social disapproval ones.

Guided by the above considerations, this study proposes to investigate concept attainment of poor premorbid schizophrenics, good premorbid schizophrenics, and normal subjects on materials portraying an adult and a child in two types of interaction, an approval interaction and a disapproval interaction. These scenes, described more fully below, vary only in the sex and dress of the depicted individual (Bruner, Goodnow, & Austin, 1956). Further, the process of concept attainment, also discussed more fully below, involves successive choices made by the subject until he formulates the rule for segregating exemplars of the concept from nonexemplars.

The following hypotheses are then indicated:

1. The poor premorbid schizophrenics, the good premorbid schizophrenics, and the normals will be significantly different in concept attainment involving social materials with the poor premorbids showing the greatest impairment and the normals the least impairment on the following variables: (*a*) the number of card choices necessary for solution; (*b*) the number of redundant card choices, that is, the card choices that provide no new information about the concept; and (*c*) total time taken for solution.

2. In the transition from concept attainment on approving interpersonal scenes to disapproving interpersonal scenes, the normals, the good premorbids, and the poor premorbids will show significantly and increasingly greater decrements on the following variables: (*a*) card choices, (*b*) redundant choices, and (*c*) total time.

METHOD

Subjects. Thirty-two male patients with the diagnosis of schizophrenia, from the Northampton, Massachusetts, State Hospital were selected, with 16 patients in the good premorbid group and the other 16 patients in the poor premorbid group. The good premorbid subjects scored 15 or below on the Phillips premorbid history scale while the poor premorbid subjects scored 20 or above on this scale. Reliability of Phillips scale scoring was .94. All schizophrenic subjects, regardless of premorbid history, were between 18 and 45 years of age, were nonhallucinating, were cooperative, and showed no other pathology, such as alcoholism, brain injury, etc. The normal subjects consisted of 16 males living and working in communities in the Northampton area, who manifested no overt personality deviancies, and who denied previous emotionally disabling illnesses.

The poor premorbid schizophrenics, the good premorbid schizophrenics, and the normals are each divided into an experimental group and a control group, making a total of six groups of eight subjects each. The six groups are matched on the basis of their average scores on the similarities and block design subtests of the WAIS; the use of these two subtests aids both in the matching for abstract ability and for IQ. A study by Maxwell (1957) indicates that these two subtests correlate .89 with full scale WAIS scores. In addition, the six subgroups are matched on age and on socioeconomic status, the latter matching accomplished by the Minnesota scale of occupations (Goodenough & Anderson, 1931).

Test Materials. The test materials comprised two 16-card sets described by Bruner et al. (1956). In one set, there is pictured, in each card, an adult figure handing a gift to a child who is smiling like the adult (approving set). In the other set, the adult figure in each card has arms clasped behind him and is frowning at a child who is looking down (disapproving set). Each card measures $4'' \times 7''$ and regardless of what set it is in, contains a combination of four attributes, that is, sex of the adult, sex of the child, type of dress of the adult, type of dress of the child. Each attribute has two values. For the sex of the adult, there are male and female, and for the dress of the adult, there are day and night dress. The same values hold true for the attributes of the child figures. The 16 cards of each set, approving and disapproving, are shown in Table 1, which also indicates the arrangement which was maintained for all subjects throughout the experiment. The type of concept to be attained in this investigation is a conjunctive concept and is defined by Bruner et al. as one where instances embraced by the concept share some common attribute or attributes.

Procedure. All the subjects were individually tested. After the completion of the two practice problems,

TABLE 1

CARD ARRANGEMENT MAINTAINED FOR ALL SUBJECTS OVER ALL PROBLEMS

FAD, MCD FAD, FCD	FAD, MCN FAD, FCN	FAN, MCD FAN, FCD	FAN, MCN FAN, FCN	MAD, MCD MAD, FCD	MAD, MCN MAD, FCN	MAN, MCD MAN, FCD	MAN, MCN MAD, FCN

Note.—F = Female, A = Adult, D = Day Clothes, M = Male, C = Child, N = Night Clothes.

both the experimental and control subjects solved the first four problems, the first half of the experiment, on the approving set of cards. In the second half, the control subjects continued on the approving set of cards for the second four problems after pausing for 1 minute during which time the approving set was gathered up and laid down once again in prescribed fashion. In their second half the experimental subjects completed the second four problems on the disapproving set. All subjects, whether experimental or control, completed the same eight problems in addition to the same two practice problems. Table 2 presents the procedure of this study.

At the beginning of each session, the following instructions were read aloud by the experimenter while the subject read silently. The instructions were reread when necessary. The completion by each subject of two practice problems further ensured that the subject understood the nature of the task.

As you see, I have sixteen cards laid out here. No two cards are exactly alike. But certain parts of each of them are like certain parts in others.

As they are arranged, you will note that in four cards there is an adult female and a boy pictured. In these four cards, the female is dressed in either night clothes or day clothes; the same with the boy, who is in either night clothes or day clothes. In the four cards below there are again adult women, but this time they are with female children, and as in the above four cards, these cards have figures with either day clothes or night clothes.

Here we have four cards again, but this time, as you can see, the adult figures are male. The children here are boys. And except for the change from female adult to male adult, these cards vary as do the other cards we have seen. And finally we have four cards where the adult male is with a female child. These cards vary as do the other cards we have seen, as you can see. Have you any questions before I continue?

Well, as I mentioned earlier, no two of these cards are exactly alike, except in terms of what the people are doing, but aspects of each card are similar to aspects of certain other cards. The aspects that we are concerned with are the sex of the figures, and their dress.

What I want you to do is to, by using these cards, attain a concept; that is, were I to ask you in what way are a wheel and a baseball alike, you would say, "they're both round," and the concept would therefore be "roundness." Well, it's the same idea here. You select cards that you think have something in common with the cards I put before you. The card that I pick will have on it pictured the concept that I have in mind, a concept like "roundness" but here it would be such a concept as "people in day clothes." To help you guess the concept, you pick cards until you get one that I tell you has on it a picture suggesting the concept that I have in mind. This would make your and my card alike in some way. When you get such a card, I will allow you to guess at the concept I have in mind. If you guess wrong—cards can be alike in more than one way—I will allow you to pick other cards, but never that same card, until you again get a card that has pictured on it the concept I have in mind and that, therefore, is pictured on the card that I have put before you. When you again get such a card, you can make another guess. The job is to keep on choosing and when indicated, guessing, until you solve the problem by giving me the correct concept. Examples of concepts are, "Boys in Day Clothes," "Males," and "People in Night Clothes."

To help make this clear, we will try two practice problems, and I will help you with them.

TABLE 2

PRESENTATION OF EXPERIMENTAL PROCEDURE

Group	N	Practice	First half	Second half
		2 Problems	4 Problems	4 Problems
Poor premorbids				
Experimental	8	Approval cards	Approval cards	Disapproval cards
Control	8	Approval cards	Approval cards	Approval cards
Good premorbids				
Experimental	8	Approval cards	Approval cards	Disapproval cards
Control	8	Approval cards	Approval cards	Approval cards
Normals				
Experimental	8	Approval cards	Approval cards	Disapproval cards
Control	8	Approval cards	Approval cards	Approval cards

Prior to starting work on the second four problems, each experimental subject was told the following:

Again with these problems we are interested only in the sex and dress of the figures.

The problems and the focus card (the first card presented by the experimenter as a positive instance containing the concept) are described below.

A. Practice Problems

1. Female Children in Day Clothes
   Focus card: Females in Day Clothes

2. Figures of the same Sex
   Focus card: Males in Night Clothes

B. Problems (First half)

1. Female Adults in Day Clothes
   Focus card: Females in Day Clothes
2. People in Night Clothes
   Focus card: Males in Night Clothes
3. Male Children in Day Clothes
   Focus card: Males in Day Clothes
4. Females
   Focus card: Females in Night Clothes

C. Problems (Second half)

5. Male Adults in Day Clothes
   Focus card: Males in Day Clothes
6. People in Day Clothes
   Focus card: Females in Day Clothes
7. Female Children in Night Clothes
   Focus card: Females in Night Clothes
8. Males
   Focus card: Males in Night Clothes

D. The dependent variables are

1. The number of cards chosen before solution, a measure of the efficiency displayed in attaining the concept.

2. The number of redundant choices. A redundant choice was a choice of a card which contained no new information.

3. Total time required before the solution of the conceptual problem.

RESULTS

The matching of the six groups is evaluated by analysis of variance. There are no significant differences in the control variables—age, abstract ability, and socioeconomic status—on which the six groups are matched (Table 3).

TABLE 3

MEANS AND STANDARD DEVIATIONS OF THE EXPERIMENTAL AND CONTROL SUBJECTS FOR AGE, ABSTRACT ABILITY, AND SOCIOECONOMIC STATUS

	Poor premorbid		Good premorbid		Normal	
	Experimental	Control	Experimental	Control	Experimental	Control
Age						
M	32.00	33.63	34.63	34.63	31.11	31.87
SD	8.80	4.43	7.19	4.58	7.96	7.98
Abstraction						
M	10.38	10.50	10.50	10.56	10.50	10.75
SD	1.55	2.20	1.45	2.74	1.85	1.49
Socioeconomic						
M	5.38	5.38	5.25	5.50	5.25	5.00
SD	1.69	2.50	1.98	1.77	1.67	1.85

Card Choices. Because there is some heterogeneity of variance in the data, a more stringent significance level, the .01 level of significance, is used instead of the .05 level with regard to the card choices, redundant choices, and total time variables.

The results of the analysis of variance dealing with the total number of card choices prior to solution reveal a significant difference when the three personality groups are compared (Table 4). Further analysis of this significant finding by the Duncan range test (Federer, 1950) indicates that the poor premorbids, good premorbids, and the normals are significantly differentiated from each other in total choices necessary for solution on the combined approval-approval and approval-

TABLE 4

ANALYSIS OF VARIANCE TABLE FOR CARD CHOICES, REDUNDANT CHOICES, AND TOTAL TIME

Source	*df*	Card choices		Redundant choices		Total time	
		SS	*F*	*SS*	*F*	*SS*	*F*
Total	95	6554		4545		1703	
Between subjects	47	4761		3450		1385	
Personality (Per)	2	2330	21.4****	1585	19.0****	784	36.4****
Pattern (Pat)	1	117	2.1	84	2.0	91	8.4**
Per × Pat	2	24	.2	26	.3	58	2.7
Error between	42	2290		1755		452	
Within subjects	48	1794		1095		318	
Halves (H)	1	26	.9	17	.1	6	1.1
H × Per	2	139	2.4	96	2.7	28	2.8
H × Pat	1	308	10.5***	170	9.6***	71	14.2***
H × Per × Pat	2	89	1.5	71	2.0	12	1.2
Error within	42	1239		742		201	

** Significant at .01 level.
*** Significant at .005 level.
**** Significant at .001 level.

TABLE 5

MEANS[a] AND STANDARD DEVIATIONS BY HALVES ARRANGED BY PATTERNS WITHIN PERSONALITY GROUPS FOR THE VARIABLES OF CARD CHOICES, REDUNDANCY, AND TOTAL TIME

	Card choices		Redundancy		Total time	
	First	Second	First	Second	First	Second
Poor premorbids						
Approval-Approval						
M	23.50	22.75	11.38	10.25	865.88	705.50
SD	9.87	8.72	8.13	8.00	335.25	225.30
Approval-Disapproval						
M	20.13	30.00	9.00	17.38	797.50	1138.36
SD	6.86	9.11	5.56	7.47	423.50	374.50
Good premorbids						
Approval-Approval						
M	16.88	13.00	6.25	3.38	644.13	405.50
SD	6.14	6.09	5.20	5.76	380.80	237.19
Approval-Disapproval						
M	16.25	20.50	6.38	8.88	695.13	1002.75
SD	5.12	8.42	3.44	7.54	243.25	350.12
Normals						
Approval-Approval						
M	12.75	10.25	2.63	1.13	167.25	155.00
SD	2.96	1.67	1.55	1.36	57.48	107.09
Approval-Disapproval						
M	12.63	12.38	2.50	2.13	223.38	206.75
SD	3.73	1.85	2.13	1.24	105.10	45.55

[a] For each mean $N = 8$.

disapproval patterns. The poor premorbids, the good premorbids, and the normals are rank ordered in choices with the poor premorbids needing the most choices to attain the concepts (Table 5).

In addition, there is a significant finding (Table 6) which indicates that the number of choices decreases from the first-half approval cards to the second-half approval cards (approval-approval pattern) while the number of choices increases from the first-half approval cards to second-half disapproval cards (approval-disapproval pattern). In order to determine whether this significant finding could be attributed to the performance of the psychiatric groups, two further analyses of variance are carried out (Table 7). One analysis of variance deals with card choices in the approval-approval pattern alone while the other analysis deals with card choices in the approval-disapproval pattern. There is a significant finding for the personality groups in both analyses. Application of the Duncan test indicates that the poor premorbids require significantly more choices than the good premorbids and the normals but that the good

TABLE 6

MEANS AND STANDARD DEVIATIONS BY HALVES ARRANGED BY PATTERNS[a] FOR THE VARIABLES OF CARD CHOICES, REDUNDANCY, AND TOTAL TIME

	Card choices		Redundancy		Total time	
	First	Second	First	Second	First	Second
Approval-Approval						
M	17.63	15.17	6.75	4.92	554.92	422.00
SD	8.02	7.92	6.89	5.02	410.00	323.00
Approval-Disapproval						
M	15.08	20.96	8.95	9.46	572.00	782.63
SD	8.80	10.01	6.08	9.56	375.00	506.00

[a] In each pattern $N = 24$.

TABLE 7

ANALYSIS OF VARIANCE TABLE FOR THE APPROVAL-APROVAL AND APPROVAL-DISAPPROVAL PATTERNS SEPARATELY CALCULATED FOR CARD CHOICES, REDUNDANT CHOICES, AND TOTAL TIME

Source	*df*	Card choices				Redundant choices				Total time			
		Approval-Approval		Approval-Disapproval		Approval-Approval		Approval-Disapproval		Approval-Approval		Approval-Disapproval	
		SS	*F*	*SS*	*F*	*SS*	*F*	*SS*	*F*	*SS*	*F*	*SS*	*F*
Between subjects	23	2332		2311		1688		1678		562		724	
Personality	2	1089	9.2***	1274	12.9****	663	6.8**	946	13.6****	315	13.4****	533	26.6****
Error	21	1243		1037		1024		731		247		190	
Within subjects	24	667		1024		385		710		88		245	
Halves (H)	1	78	2.8	257	8.0***	40	2.5	147	7.0***	24	9.3	54	6.2*
H × Per	2	14	.3	98	3.1	7	.2	121	2.9	9	1.7	21	1.3
Error	21	576		671		337		441		54		170	

* Significant at the .05 level.
** Significant at the .01 level.
*** Significant at the .005 level.
**** Significant at the .001 level.

premorbids are not different from the normals in the approval-approval pattern. For the approval-disapproval pattern, the poor premorbids, the good premorbids, and the normals are significantly rank ordered with the poor premorbids requiring the greatest number of choices before attaining the concept. But in the approval-approval pattern and in the approval-disapproval pattern as analyzed separately (Table 7), there are no significant interactions. The three personality groups, hence, suffer about the same decrement in choices, first in the transition from one approval half to the other approval half of the approval-approval pattern, and second in the transition from the approval half to the disapproval half of the approval-disapproval pattern. Finally, no significant finding results from the comparison of the differences between card choices of poor premorbids, good premorbids, and normals, respectively, on the first and second halves for each approval-approval and approval-disapproval pattern.

Redundant Choices. The analysis of variance indicates that there is a significant difference for the three personality groups in the number of redundant choices. (A redundant choice is a chosen card that adds no new information about the correct concept, Tables 4 and 5.) The Duncan test, applied to this finding, indicates that the poor premorbids, good premorbids, and normals are rank ordered with the poor premorbids having significantly more redundant choices than the good premorbids who in turn have significantly more redundant choices than the normals.

Another significant finding (Tables 4 and 6) indicates that the number of redundant choices decreases from the first-half approval cards to the second-half approval cards (approval-approval pattern) while the number of redundant choices increases from the first-half approval cards to the second-half disapproval cards (approval-disapproval pattern). In order to determine whether this significant finding could be attributed to the performance of the psychiatric groups, two further analyses of variance are calculated (Table 7). One analysis is concerned with redundant choices in the approval-approval pattern alone while the other analysis is concerned with redundant choices in the approval-disapproval pattern. The results indicate significant differences for the personality groups in the approval-approval pattern and in the approval-disapproval pattern. Using the Duncan test, it is determined that the poor premorbids have significantly more redundant choices than the good premorbids and the normals on redundant choices but that there is no difference between the good premorbids and the normals in the approval-approval pattern. For the approval-disapproval pattern, all three personality groups are significantly different with the poor premorbids having the greatest number of redundant choices. But in the ap-

proval-approval pattern and in the approval-disapproval pattern as analyzed separately (Table 7), there are no significant interactions. The three personality groups, hence, suffer about the same decrement in redundant choices, first in the transition from one approval half to the other approval half of the approval-approval pattern, and second in the transition from the approval half to the disapproval half of the approval-disapproval pattern.

Finally, no significant finding results when there is a comparison of the differences between redundant card choices of poor premorbids, good premorbids, and normals, respectively, on the first and second halves for each approval-approval and approval-disapproval pattern.

Total Time. The analysis of variance shows that for the personality groups there is a significant finding on the variable of total time (Table 4). And on the basis of the Duncan findings for this measure, we can state that the poor premorbids, the good premorbids, and the normals are significantly different from each other in total time, with the poor premorbids taking the most time and the normals the least time.

Another significant finding (Table 6) reveals that total time decreases from the first-half approval cards to the second-half approval cards (approval-approval pattern) while total time increases from the first-half approval cards to the second-half disapproval cards (approval-disapproval pattern). In order to determine whether this significant interaction could be attributable to the clinical groups, two further analyses of variance are calculated (Table 7). One analysis is concerned with total time in the approval-approval pattern alone while the other is concerned with total time in the approval-disapproval pattern. The results indicate that there are significant findings for the personality groups in the approval-approval pattern as well as in the approval-disapproval pattern. Duncan tests reveal that the poor premorbids, good premorbids, and the normals are significantly different on the approval-approval pattern with regard to total time, with the poor premorbids requiring the most time and the normals requiring the least time. On the approval-disapproval pattern, the normals take significantly less time than the good premorbids and poor premorbids; the good premorbids and the poor premorbids are not significantly different on total time. But in the approval-approval pattern and in the approval-disapproval pattern as analyzed separately (Table 7), there are no significant interactions. The three personality groups, hence, suffer about the same decrement in total time, first in the transition from one approval half to the other approval half of the approval-approval pattern, and second in the transition from the approval half to the disapproval half of the approval-disapproval pattern.

Finally, no significant findings appear when comparisons are made of the differences between redundant card choices of poor premorbids, good premorbids, and normals, respectively, on the first and second halves for each approval-approval and approval-disapproval pattern.

DISCUSSION

Performance on Combined Approval and Disapproval Social Materials. It appears that the first hypothesis of significant differences between poor premorbids, good premorbids, and normals in concept attainment on social materials is supported by the results. On the three variables—card choices, redundant choices, and total time—the poor premorbids, the good premorbids, and the normals are significantly rank ordered for their total performance on both the approval-approval and approval-disapproval patterns, with the poor premorbids showing the poorest performance.

At this point, it should be recalled that the groups have been matched on abstract ability as measured by two Wechsler Intelligence Scale subtests. The first subtest, similarities, evaluates verbal conceptual ability (Rapaport, Gill, & Schafer, 1945) and the other, the block designs, is considered as well as a test of concept formation (p. 273, footnote) or a test of abstract ability (Goldstein & Scheerer, 1941). Despite this matched ability to perform on formal conceptualization tests, there is the finding that poor premorbids are significantly inferior than good premorbids who in turn are significantly inferior than nor-

mals in concept attainment when the stimulus materials involve social interaction. The greater the social disarticulation, as manifested by poor premorbids, the greater is the selective conceptual impairment.

Here, conceptual deficit seems to be employed by the schizophrenics against threat issuing from social interaction, and abstract operations concerning the relevant properties of the stimulus materials seem to be eschewed (Cameron, 1947; Goldstein, 1959; Sullivan, 1956). In this type of deficit, there appears to be manifested what Werner (1948) defines as syncretic activity; emotional or need factors are fused with abstract functions rather than being subordinated to these functions. On this basis, one may infer that the conceptual operations of the schizophrenics may reflect a more immature developmental level when dealing with social materials.

From Bruner's (1957), viewpoint on perceptual unreadiness, schizophrenics appear less ready than normals to attain concepts involving social materials since they require more card choices, more redundant card choices, and more time. Possibly, this greater unreadiness of schizophrenics results from their more stunted capacities in classifying and relating events involving social materials as well as from their needs and defenses which cause classes of feared events to have narrower limits of acceptance for quite relevant information necessary to concept attainment.

A close look at the nature of the concept attainment task reveals that the subject has to determine the attribute value or values—in each case involving people—that define the concept. After the experimenter advises the subject that a chosen card is an exemplar of the concept, the subject is free to state in what way the focus card and the chosen exemplar are alike. This process requires the subject to discriminate the different attribute values possessed by each card, to identify the attribute value or values that they share, and to verbalize the attribute value or values defining the concept, taking into account the information provided by the experimenter about previous choices. Possibly, because the task involves cards representing social interaction, and people as the defining attributes of the correct concept, the process of arriving at the correct concept is interfered with by the stunted capacities and fearful anticipations of the schizophrenics.

Performance on the Approval Materials as Contrasted with the Disapproval Materials. The second hypothesis of increasingly greater decrements in conceptual performance on disapproving interpersonal materials as contrasted to approving interpersonal materials for normals, good premorbids, and poor premorbids, respectively, is not supported. The implications that we have drawn from the theories of Goldstein (1959), Cameron (1947), and Sullivan (1956), which indicate that schizophrenics have a learned vulnerability to cues of interpersonal threat such as censure and disapproval, are not confirmed. These results parallel those of Hellman and Kates (1961) indicating that experimenter invoked censure does not lead to increasingly greater significant conceptual decrements in normals, good premorbids, and poor premorbids, respectively. The findings are not consistent with the analysis made by Rodnick and Garmezy (1957) that the poor premorbids in Dunn's (1954) study are significantly different than the normals in the visual discrimination tasks involving a disapproving parent-child scene.

The present study indicates that all groups show deficit behavior on social materials suggesting censure, with the psychiatric and the normal groups showing about the same relative decrements on the disapproval cards. We find that normals fail to improve their conceptual performance as much under conditions of disapproval as do their own controls under conditions of approval. Vulnerability to censure may be by no means a behavioral characteristic of the schizophrenic or of the poor premorbid schizophrenic alone, and further studies of schizophrenic vulnerability must include normal control and normal experimental groups.

It is possible that the schizophrenics, and particularly the poor premorbids, are more vulnerable to censure than the normals, but that such vulnerability depends upon cumulative effects of exposure to censure in order to be manifested as gross behavioral deficit. A greater number of problems using social disapproval materials may have this cumulative

effect. Finally, the social disapproval materials portray an adult male as well as an adult female in a disapproving relation to a child. Since there may be differential relationships to mothers and fathers for poor and good premorbids, respectively (Rodnick & Garmezy, 1957), it is possible that potentially significant decrements in concept attainment are obscured when combined father related censure and mother related censure stimulus materials are employed.

SUMMARY

This study compares concept attainment of poor premorbid schizophrenics, good premorbid schizophrenics, and normals on social materials involving interpersonal approval and disapproval. Six matched groups of eight subjects each constitute an experimental and a control group of poor premorbids, an experimental and a control group of good premorbids, and an experimental and a control group of normals. The experimental groups solve four problems on 16 cards portraying an approving scene between an adult and a child and four problems on 16 cards portraying a disapproving scene between an adult and a child. The control groups solve four problems on the 16 cards portraying the approving scene and four problems on the same social approval cards again. The approving cards and the disapproving cards differ only in their portrayal of approval or disapproval as the social interaction between the adult and the child.

The dependent variables are total number of card choices to solution; number of redundant card choices, a measure of card choices that provide no new information; and total time to correct solution.

The first hypothesis is that poor premorbid schizophrenics, good premorbid schizophrenics, and normals differ significantly in concept attainment on social materials, with the poor premorbids manifesting the most severely impaired performance, the good premorbids manifesting intermediate performance, and the normals the best performance. The second hypothesis indicates that the poor premorbids will show the greatest decrement when their concept attainment performance on the disapproval materials is compared with their performance on approval materials, that the good premorbids will show the intermediate decrement, and that the normals will show the smallest decrement, and that these decrements will be significantly different.

As predicted in the first hypothesis, all three groups evidence significant differences in concept attainment on the three response variables, number of card choices, redundant choices, and total time. Despite their being matched in conceptual ability on formal tasks with the normals, the schizophrenic groups manifest selective impairment in concept attainment on social materials. Apparently conceptual deficit may be a defense against the social environment perceived as hostile and threatening by the schizophrenics. This defense is seen as resulting in greater unreadiness on the part of the schizophrenics than normals to attain concepts involving social materials. And the greater the degree of social disarticulation as seen in the poor premorbids, the greater is the inferiority of concept attainment performance.

With respect to the second hypothesis, the groups do not differ significantly on the three response variables, when their respective performances on the approval cards are contrasted with their performances on the disapproval cards. A greater number of concept attainment problems, it is thought, may be essential to induce more gross conceptual deficit in the clinical groups.

REFERENCES

BLEKE, R. C. Reward and punishment as determiners of reminiscence effects in schizophrenic and normal subjects. *J. Pers.*, 1955, **23**, 479–498.

BRUNER, J. S. On perceptual readiness. *Psychol. Rev.*, 1957, **64**, 123–152.

BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. *A study of thinking*. New York: Wiley, 1956.

CAMERON, N. *The psychology of behavior disorders: A biosocial interpretation*. New York: Houghton Mifflin, 1947.

DAVIS, R. H., & HARRINGTON, R. W. The effects of stimulus class on the problem solving of schizophrenics and normals. *J. abnorm. soc. Psychol.*, 1957, **54**, 126–128.

DUNN, W. L. Visual discrimination of schizophrenic subjects as a function of stimulus meaning. *J. Pers.*, 1954, **23**, 48–64.

FEDERER, W. T. *Experimental design*. New York: Macmillan, 1950.

GOLDSTEIN, K. Concerning the concreteness in schizophrenia. *J. abnorm. soc. Psychol.*, 1959, **56**, 146–148.

GOLDSTEIN, K., & SCHEERER, M. Abstract and concrete behavior. *Psychol. Monogr.*, 1941, **53**(2, Whole No. 239).

GOODENOUGH, F. L., & ANDERSON, J. E. *Experimental child study*. New York: Century, 1931.

HEATH, D. H. Individual anxiety thresholds and their effect on intellectual performance. *J. abnorm. soc. Psychol.*, 1956, **52**, 403–408.

HELLMAN, B., & KATES, S. L. Conceptual performance in schizophrenics as a function of premorbid adjustment level and verbal censure. Paper presented at American Psychological Association, New York, August 1961.

MAXWELL, EILEEN. Validities of abbreviated WAIS scales. *J. consult. Psychol.*, 1957, **21**, 121–126.

PHILLIPS, L. Case history data prognosis in schizophrenia. *J. nerv. ment. Dis.*, 1953, **117**, 515–525.

RAPAPORT, D., GILL, M., & SCHAFER, R. *Diagnostic psychological testing*. Vol. 1. Chicago: Year Book Publishers, 1945.

RODNICK, E. H., & GARMEZY, N. An experimental approach to the study of motivation in schizophrenia. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1957*. Lincoln: Univer. Nebraska Press, 1957. Pp. 109–184.

SULLIVAN, H. S. *Clinical studies in psychiatry*. New York: Norton, 1956.

WEBB, N. W. Conceptual ability of schizophrenics as a function of threat of failure. *J. abnorm. soc. Psychol.*, 1955, **50**, 221–224.

WERNER, H. *Comparative psychology of mental development*. (Rev. ed.) New York: International Univer. Press, 1948.

WHITEMAN, M. The performance of schizophrenics on social concepts. *J. abnorm. soc. Psychol.*, 1954, **49**, 266–271.

(Received September 5, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 365–375

THE INFLUENCE OF POSITIVE INTERPERSONAL RELATIONS ON ADJUSTMENT AND EFFECTIVENESS IN RIFLE TEAMS [1]

JOSEPH E. McGRATH

University of Illinois

The present study is part of a program of research dealing with quasi-therapeutic relations in small work groups. By quasi-therapeutic relations we mean supportive interpersonal relations among members of work groups which, like the therapeutic situation, provide an environment that facilitates adjustment of the group member. This study was conducted in the context of a rifle team tournament and involved the experimental assembly of teams on the basis of prior interpersonal relations. How these interpersonal relations affected individual adjustment and marksmanship performance is the central focus of the study.

Background

In recent years there has been considerable speculation about the value which psychotherapeutic services have in work relations. In particular, Rogers and his co-workers (1951), Gordon (1951), and the Tavistock Institute in London (e.g., Rice, 1958), have urged the application of psychotherapeutic techniques as a means of improving work performance. The basic assumption has been that secure and anxiety free interpersonal relations will not only aid the individual's adjustment but will also help him to perform with increased efficiency. The strong appeal of this argument has led organizations ranging from Indian textile mills (Rice, 1958) to baseball teams to avail themselves of essentially psychotherapeutic services. More recently, Fiedler, Hutchins, and Dodge (1959) have suggested that it might be possible to reduce the anxiety and to improve the work efficiency of group members by judiciously placing them with others who are basically accepting.

There seems to be considerable evidence that positive interpersonal perceptions within a group lead to good interpersonal relations and facilitate individual adjustment. For instance, Fiedler et al. (1959) studied interpersonal perceptions in a variety of groups, including university dormitory groups, Army tank crews, and antiaircraft crews. Results showed that individuals who perceived others as similar, or who were seen as similar by members of their group, tended to be better adjusted and to improve on personal adjustment indices over relatively short time spans. From these results, the authors postulated that favorable interpersonal perceptions may lead to "quasi-therapeutic" relations within the group, which in turn facilitate individual adjustment. A subsequent investigation by Alexander and Drucker (1960) demonstrated that experimentally induced perceptions of similarity aided the adjustment of group members, while individuals who were seen as dissimilar tended to adjust less well to the group.

However, there is some reason to question the assumption that positive interpersonal relations are necessarily conducive to task efficiency. Fiedler and his co-workers (1958) have shown that psychologically close interpersonal relations between leader and followers in task groups may be detrimental to team performance. Furthermore, while the Fiedler et al. (1959) study suggested that good interpersonal relations aid both individual adjustment and task effectiveness, the findings were based primarily on relationships between interpersonal perceptions and indices of adjustment of the self-report type. A later study by Hutchins and Fiedler (1960) reported slight negative correlations between individual adjustment and the effectiveness

[1] This investigation was supported by Contract Number DA-49-193-MD-2060, "Interpersonal Perception and the Psychological Adjustment of Group Members," between the Office of The Surgeon General, Department of the Army, and the University of Illinois (Fred E. Fiedler, principal investigator).

of antiaircraft artillery crews. A recent review of small group research studies by McGrath and his associates (Altman & McGrath, 1959; Altman & Terauds, 1960; McGrath, 1957; Terauds, Altman, & McGrath, 1960) lent further support to the argument that good interpersonal relations within work groups do not necessarily lead to increased individual task effectiveness. Examination of 250 small group research studies showed very few instances of statistical tests, and no instances of significant relationships, between objectively scored indices of individual performance and measures indicative of good interpersonal relationships.

Thus, while it seems fairly well established that positive interpersonal relationships within a group aid the adjustment of group members, it is by no means clear that such relationships have a correspondingly good effect upon the task performance of individuals. The present study was designed as a specific test of the assumption that a positive interpersonal climate within a small work group will lead to both better adjustment and better individual task performance by its members.

General Approach

Two features of the study are worthy of note. First, the experimental task, rifle marksmanship, offers several advantages. It is intrinsically interesting to young men; it yields objective and highly reliable individual performance scores; and, like tremor and concentration measures, marksmanship is likely to be affected by tenseness and therefore should be sensitive to changes in individual adjustment. Rifle marksmanship performance, however, is essentially a series of independent individual performances. Little or no coordination among team members is required. Thus, it may be less sensitive to differences in interpersonal relations than other more interactive group tasks.

The second notable feature of the study is that teams were experimentally assembled on the basis of the participants' prior interpersonal relations. This procedure enabled us to carry our research on quasi-therapeutic relations (Alexander & Drucker, 1960; Fiedler et al., 1959; Myers, 1961) the next logical step forward. It also enabled us to test a key methodological question in that program. The Fiedler et al. postulation of "quasi-therapeutic" relations within work groups does not indicate whether such relations arise from perceptual or behavioral characteristics of group members. If A sees Teammate B as a warm, supportive individual, is it because B acts toward A in certain ways or because A tends to perceive such supportiveness in his social environment? The answer to this question was of methodological and substantive significance to the research program.

The study, thus, had two objectives: to assess the effects of positive interpersonal relations on individual adjustment and task effectiveness; and to determine whether the formation of positive interpersonal relations in groups depends primarily on the perceptions or the behavior of group members.

PROCEDURE

The present study took advantage of an ongoing experiment in which 60 three-man teams participated in a 5-week rifle marksmanship tournament. This experiment by Myers (1961) investigated the effects which competition and team success had on interpersonal perceptions and on adjustment in three-man rifle teams.[2] The tournament participants were 180 Army ROTC volunteers of varying levels of marksmanship ability, who were organized into three-man teams. Measures of individual adjustment, interpersonal perception, and individual and team marksmanship were obtained during and at the end of the marksmanship tournament.

Responses to a sociometric nomination question and four behavior description items at the end of the regular tournament were used to construct an index of interpersonal relations between teammates. The sociometric nomination question asked which teammate(s), if either, "helped you stay calm and relaxed while shooting." The four behavior description items were: "He was a warm person, easy to talk to"; "He was sort of standoffish" (scored in the reverse direction); "He listed attentively to others"; and, "He was disruptive to the group" (scored in the reverse direction). Participants rated each of their teammates on five-point rating scales. For purposes of team assembly, the nominations and ratings were combined into a single index which gave the sociometric nomination a weight equal to the sum of the four behavior description ratings.

Each man was assigned two scores: the total interpersonal relations scores which he *gave* to his two tournament teammates; and the total interpersonal

[2] We wish to thank Albert E. Myers for cooperation in use of the subject population, and to gratefully acknowledge the assistance of James W. Julian, and Scott E. Krueger, all of the Group Effectiveness Research Laboratory, University of Illinois, in the data collection phase of this study. We are especially appreciative of the cooperation which we received from Clair Worthy, Commander of the University of Illinois Army ROTC unit, and his staff. In particular, D. McClelland, W. Hooper, and M. Snoke helped to coordinate the project. E. F. Flanders served as consultant and range safety officer.

relations score which he *received* from his two teammates. The first is a "perceptual index," indicating the extent to which the individual saw others as warm, supportive, and accepting. The second is a "behavior index," which reflects the extent to which others saw him as exhibiting such positive interpersonal behaviors.

New teams were composed by putting together sets of three men who had high scores on the perceptual index but low scores on the behavioral index; sets of three men who had low scores on the perceptual index but high scores on the behavioral index; and sets of three men who had intermediate scores on both indices. During the following week (Week 6 of the experiment) these new teams participated in 1-hour, six-trial, handicap marksmanship contest for a money prize, and then completed the same battery of questionnaires that had been utilized in the first phases of the rifle team experiment.

The use of intrateam rankings (rather than absolute scores) for reassembling teams tended to reduce the average score differences between the three sets of experimental teams. Therefore, only data from the two extreme conditions—high scorers on the behavioral index, and high scorers on the perceptual index—were utilized in the analyses, although the intermediate scoring teams were permitted to participate in the marksmanship contest.

Thirty-five teams (105 participants) were available for the Week 6 experimental sessions. Complete data were obtained for 24 individuals (10 teams) scoring high on the behavior index and 29 individuals (11 teams) scoring high on the perceptual index.[3] All analyses presented in this paper are based on data from these 53 individuals at completion of the regular tournament (Week 5) and in the reassembled teams (Week 6).

Measures

Marksmanship. In both the regular tournament and the reassembled teams a "trial" consisted of a three-round shot pattern by each member of the team. Individual marksmanship scores for a given trial were obtained by measuring the distance between the two most separated rounds of the shot pattern in ⅛-inch units. Individual average marksmanship scores were computed for each participant for all trials of Week 5 and for all trials of Week 6.

Interpersonal Perception Tests (IPT). The basic form of the interpersonal perception tests (IPT) used in this study was a 24-item questionnaire similar to the Semantic Differential. Each item consisted of an eight-point scale whose ends were defined by a pair of polar opposites; e.g., Talkative-Not talkative. Individuals were asked to describe themselves or other persons by checking one of the eight spaces on each of the 24 items. IPT responses were obtained for self at the end of the tournament (Week 5) and for each teammate in the tournament teams (Week 5) and in the reassembled teams (Week 6).

Esteem scores were derived from IPT responses by assigning maximum score (8 points) to the most favorable end of each item and minimum score (1 point) to the least favorable end, and then summing scores for all 24 items. The following esteem scores were used: Self-Esteem, Esteem for Teammates (the average of the esteem given to the individual's two teammates), and Esteem Received from Teammates (the average of the esteem scores which the individual received from his two teammates).

Postmeeting Questionnaire. This instrument consisted of eight five-point rating scales and four nomination questions. It was administered each week of the study. Several measures were derived from responses to these questions: ***Social Adjustment,*** which was operationally defined as the sum of responses to five rating items asking to what extent the individual enjoyed being a member of his team and felt accepted and comfortable in it; ***Task Satisfaction,*** which was operationally defined as the response to a rating item asking the individual how satisfied he was with his own task performance; ***Extrapunitiveness toward Teammates,*** which was based on the response to a question asking who or what was to blame when the team did poorly; and ***Positive Interpersonal Nominations,*** which were the nominations an individual received from his teammates on a question asking which teammate, if either, helped him to remain calm and relaxed while shooting. Responses to the latter measure at the end of the tournament were used as part of the basis for assembly of teams in Week 6. Nominations received in the Week 6 teams were used as one measure of the effect of reassembly.

Behavior Description Questionnaire. This instrument consisted of 25 five-point rating scales asking individuals to describe the behavior of each of their teammates. It was administered in both Week 5 and Week 6. As indicated above, responses to four of these items in Week 5 were used as part of the index of positive interpersonal relations for assembly of Week 6 teams. Responses in Week 6 to those same four rating items were used to obtain two scores for each participant: a ***behavioral index of interpersonal relations,*** which was operationally defined as the sum of the responses which the individual received from his two teammates on these items; and a ***perceptual index of interpersonal relations,*** which was operationally defined as the sum of the responses which the individual gave to his two teammates on these four items. In addition, mean ratings given and mean ratings received on each of the 25 items of the Behavior Description

[3] A total of 35 teams (105 men) was available for this experiment. Elimination of the intermediate condition (12 teams, 36 men), absenteeism in the sixth week of the study, plus 2 individuals who participated in the marksmanship task but did not complete questionnaires, accounts for the attrition down to 53 subjects.

Questionnaire were computed for members of the two experimental conditions.

RESULTS

Perceptual and Behavioral Indices of Positive Interpersonal Relations

The experimental assembly procedures produced two sets of rifle teams representing two contrasting patterns of interpersonal relations. Members of one set of teams were high on the perceptual index but low on the behavioral index of positive interpersonal relations. That is, they had seen former teammates as warm and supportive but had been judged less favorably by those teammates. Members of the other set of teams showed the obverse pattern; they had been judged by former teammates as warm and supportive, but had perceived their former teammates less favorably. The first question for analysis was which of the two groups should be considered as representing a positive interpersonal relations condition.

This question was attacked by comparing the relative consistency of ratings *given* to former and new teammates versus ratings *received* from former and new teammates. If group members who are perceived favorably by one set of teammates are also perceived favorably by new teammates, then there is a basis for holding that those judgments are a function of the interpersonal behavior of the persons so rated. On the other hand, if group members who perceive one set of teammates favorably also perceive new teammates favorably, then there is a basis for inferring that those judgments are a function of the perceptual tendencies of the persons making the ratings.

Three sets of data were used to make these comparisons: responses to the interpersonal nomination question, responses to the four interpersonal relations rating items, and responses to all 25 items of the Behavior Description Questionnaire. The first two sets of data were utilized because they paralleled, for the reassembled teams, the measures which formed the basis for team reassembly. The third set of data was used to determine whether responses to the four rating items selected for team reassembly were similar to responses to other types of behavior ratings.

The distribution of choices on the interpersonal nomination question for the reassembled teams is given in Table 1. Thirteen of the 29 members of the group which had scored high on the perceptual index (who had received few but who had given many interpersonal nominations to former teammates) received such nominations from their new teammates. In comparison, only 4 of the 24 members of the group who had scored high on the behavioral index (who had received many but who had given few interpersonal nominations to former teammates) received interpersonal nominations from their new teammates. The chi square for this distribution is significant ($p < .05$). Thus, the tendency to give positive interpersonal nominations to teammates was more consistent over team situations than the tendency to receive positive nominations from teammates.

The two experimental conditions did not differ significantly on mean scores for the four interpersonal relations items, for either the perceptual or behavioral indices. However, individuals who had scored high on the perceptual index received higher mean scores (hence, more favorable ratings) than persons who had scored high on the behavioral index for 19 of the 25 items of the Behavior Description Questionnaire. They also gave more favorable ratings on 20 of the 25 items.

TABLE 1

POSITIVE INTERPERSONAL NOMINATIONS RECEIVED IN THE REASSEMBLED TEAMS FOR THE TWO EXPERIMENTAL GROUPS

	Received positive interpersonal nominations in reassembled teams	Did not receive positive interpersonal nominations in reassembled teams	Total
Members of the Experimental group who *gave* positive interpersonal nominations to regular tournament teammates	13	16	29
Members of the Experimental group who *received* positive interpersonal nominations from regular tournament teammates	4	20	24

Note.—$\chi^2 = 4.76$, $df = 1$, $p < .05$.

Both of these distributions departed significantly from chance expectation ($p < .05$).

These results support the conclusion that perceptual aspects outweighed behavioral aspects in determining the initial interpersonal relations within the reassembled teams. Thus, while the recipient of positive interpersonal judgments may or may not be in fact ". . . a warm person . . . ," who ". . . listens attentively . . ." it appears that persons who give such positive judgments are prone to see warmth and supportiveness in their social environment.

In view of these findings, the set of teams which was composed of men who had given positive interpersonal nominations and ratings to their former teammates were designated as the Positive Interpersonal Relations (PR) group. Teams composed of men who had given low positive interpersonal relations scores to their tournament teammates were designated as the Nonpositive Interpersonal Relations (NR) group.

Marksmanship

Marksmanship scores for members of the two experimental conditions prior to reassembly (Week 5) and in the reassembled teams (Week 6) are shown in Table 2 and Figure 1. The two groups did not differ significantly in marksmanship at the end of the regular tournament. In Week 6, however, the NR group had significantly better marksmanship scores than the PR group ($p < .01$). Furthermore, the NR group showed significant improvement in marksmanship from Week 5 to Week 6 ($p < .01$), while the PR group did not improve. Marksmanship scores were also computed for the last three trials of Week 6, when the pressure of the competition was presumably at its highest and when the effects of interpersonal relations in the new teams had time to become manifest. Here an even larger difference was obtained between the two experimental conditions ($p < .01$).

These results clearly call into question the assumption that positive interpersonal relations within a work group lead to increased task effectiveness. Members of the PR group apparently had more favorable interpersonal relations—at least from the subjective viewpoints of the members—than did members of the NR group. Yet members of the NR group improved in marksmanship in the reassembled teams while members of the PR group did not improve. Since the two groups did not differ significantly in marksmanship scores while in their original tournament teams, the data suggest that assembly of the new teams somehow facilitated the marksmanship performance of members of the NR group.

TABLE 2

MEAN MARKSMANSHIP SCORES FOR THE TWO EXPERIMENTAL CONDITIONS[a]

	Nonpositive Interpersonal Relations group	Positive Interpersonal Relations group	Difference
Mean tournament marksmanship score (Week 5)	7.79	8.52	−.73
Mean marksmanship score in reassembled teams (Week 6)	6.46	8.49	−2.03**
Difference (Week 5 to Week 6)	−1.33**	−0.03	
Mean marksmanship score, last three trials for reassembled teams (Week 6)	5.95	8.19	−2.24**

[a] Marksmanship scores were computed on the basis of scatter of shot pattern. Thus, the lower the score, the better the marksmanship performance.
** $p < .01$.

Adjustment Measures

Mean scores for members of the two experimental conditions did not differ significantly on any of a number of self-report measures of adjustment and interpersonal perception, either before or after team reassembly, nor was there evidence for gross relationship between positive interpersonal relations and adjustment. Therefore, to understand better the obtained differences in marksmanship between the two experimental conditions, we turned to correlational analyses of the pattern of relationships within each of the two experimental groups.

Correlates of Performance and Adjustment

The correlational analyses were designed to discover why reassembly into new teams had a facilitative effect upon the task performance

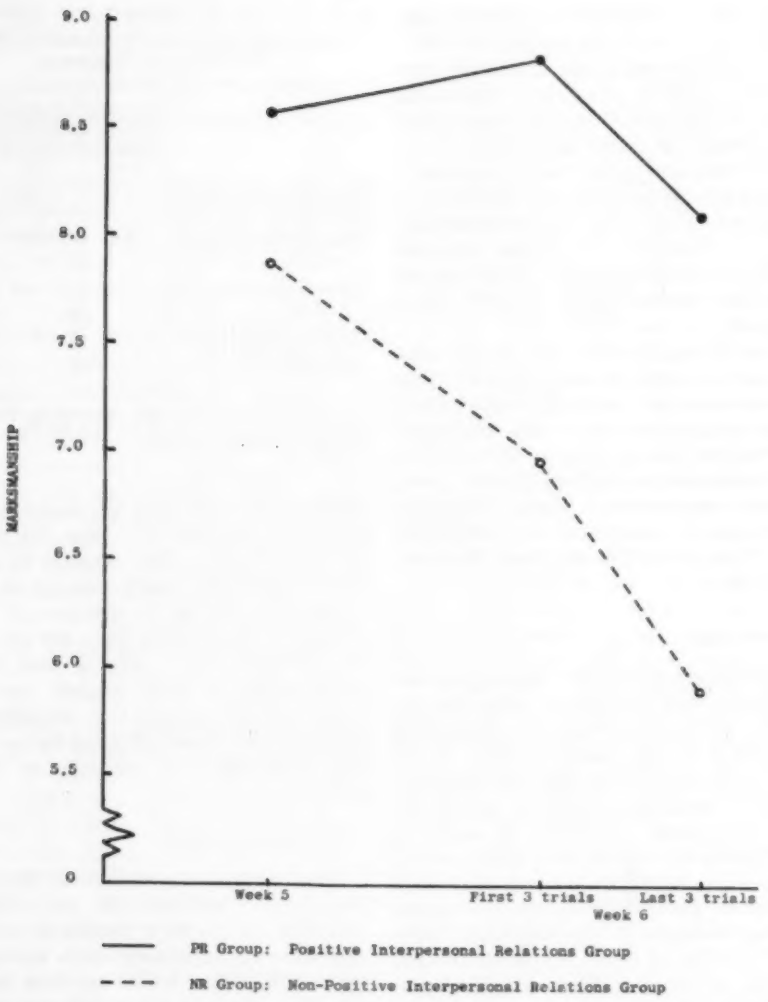

FIG. 1. Average marksmanship scores for the two experimental conditions. (Marksmanship is measured in terms of average scatter of three-round shot patterns, in ⅛-inch units.)

of members in the NR group while apparently having a deleterious effect upon their interpersonal relations. It is one thing to postulate that good interpersonal relations may or may not lead to effective task performance. It is quite another matter, theoretically, to find that *removal* of supportive teammates leads to an improvement in task performance.

We formulated the general a posteriori hypothesis that individuals in the PR group may have had an entirely different orientation toward the team-task situation than did members of the NR group. Specifically, we hypothesized that members of the PR group focused their attention and energies primarily on the social or interpersonal aspects of the team interaction, while mem-

TABLE 3

CORRELATIONS WITH MARKSMANSHIP, SOCIAL ADJUSTMENT, AND TASK SATISFACTION FOR THE TWO EXPERIMENTAL CONDITIONS

	Nonpositive Interpersonal Relations group (N = 24)			Positive Interpersonal Relations group (N = 29)		
	Marksmanship[a]	Social Adjustment	Task Satisfaction	Marksmanship[a]	Social Adjustment	Task Satisfaction
Social Adjustment	−.47*	—	—	.05	—	—
Task Satisfaction	−.39	.45*	—	−.20	.40*	—
Self-Esteem	−.04	−.01	−.21	.10	.48**	.16
Esteem for Teammates	.08	.21	.32	.07	.63**	.08
Esteem Received from Teammates	−.20	.15	.25	−.45*	.04	.06
Extrapunitiveness toward Teammates	.05	.00	.47*	−.15	−.30	.07
Interpersonal relations ratings given to teammates	.18	.55**	.27	.01	.62**	.04
Interpersonal relations ratings received from teammates	.21	.01	.34	−.27	−.06	.19

[a] Marksmanship is scored in terms of scatter of shot pattern. Thus, a low score means good marksmanship.
* $p < .05$.
** $p < .01$.

bers of the NR group concentrated on their own task performance to the neglect of interpersonal relations.

If this hypothesis were true, then adjustment to the team situation should be a function of the individual's own task performance in the NR group but not in the PR group. In contrast, the individual's adjustment to the team situation for members of the PR group should be a function of the individual's social success—how well he liked, and was liked by, his teammates.

This hypothesis was tested by examining the pattern of correlations for performance, adjustment, and interpersonal perception measures within the two experimental conditions (Tables 3 and 4). Diagrams of the pattern of significant relationships for the two experimental conditions are shown in Figures 2 and 3.

Marksmanship. In the NR group Marksmanship correlated significantly with both Social Adjustment and Task Satisfaction, which are considered two measures of the individual's adjustment to the team situation. Neither of these relationships held for the PR group. Instead, in the latter group, an individual's marksmanship was related to the favorableness of his teammates' responses to him (esteem received and behavior ratings received).

Social Adjustment and Task Satisfaction. Social Adjustment scores and Task Satisfaction were significantly correlated with each other for both groups. In addition, Social Adjustment was significantly related to perceiving teammates as supportive and accepting and to giving favorable behavior ratings to teammates in both experimental groups. However, in the PR group, Social Adjustment was also related to Self-Esteem and Esteem for Teammates, while Task Satisfaction was

TABLE 4

PROPORTIONS OF BEHAVIOR RATING ITEMS WHICH HAD POSITIVE CORRELATIONS WITH MARKSMANSHIP, SOCIAL ADJUSTMENT, AND TASK SATISFACTION FOR THE TWO EXPERIMENTAL CONDITIONS

	Percentage positive correlations for behavior ratings given to teammates	Percentage positive correlations for behavior ratings received from teammates
Nonpositive Interpersonal Relations group		
Marksmanship[a]	32	68
Social Adjustment	76*	44
Task Satisfaction	60	68
Positive Interpersonal Relations group		
Marksmanship[a]	52	8**
Social Adjustment	100**	72*
Task Satisfaction	44	72*

[a] Marksmanship is scored in terms of scatter of shot pattern. Thus, a low score means good marksmanship.
* $p < .05$.
** $p < .01$.

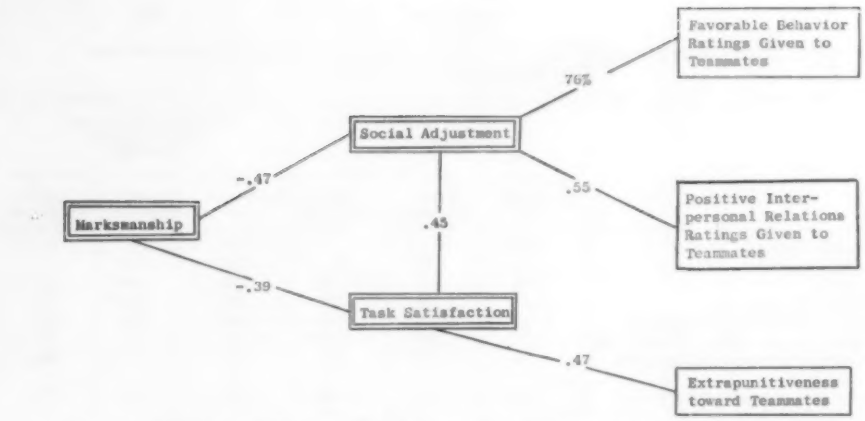

FIG. 2. Diagram of significant relationships for Nonpositive Interpersonal Relations group. (Marksmanship is scored in terms of scatter of shot pattern. Therefore, a low score indicates good marksmanship, and a negative correlation indicates that good marksmanship is associated with a high score on the other variable. Favorable Behavior Ratings Given to Teammates are percentages of 25 BDQ items which correlated in the positive direction with the variable indicated.)

related to a tendency to receive favorable behavior ratings from teammates. In the NR group, Task Satisfaction was related to a tendency to be extrapunitive toward teammates.

These correlations lend support to the hypothesis that members of the NR group were concerned more with their own task success than with interpersonal relations. In the NR group, both Social Adjustment and Task

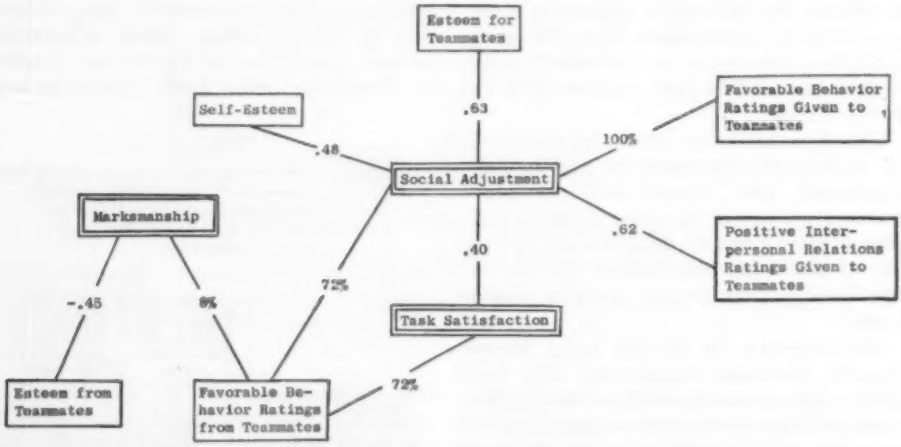

FIG. 3. Diagram of significant relationships for Positive Interpersonal Relations group. (Marksmanship is scored in terms of scatter of shot pattern. Therefore, a low score indicates good marksmanship, and a negative correlation indicates that good marksmanship is associated with a high score on the other variable. Favorable Behavior Ratings Given to Teammates are percentages of 25 BDQ items which correlated in the positive direction with the variable indicated.)

Satisfaction were related to—and perhaps derived from—success on the task. Those who shot well were satisfied with their own task performance and felt accepted by the group (although they did not actually receive more favorable evaluations by teammates). Those who shot poorly responded more negatively to the social situation. In contrast, marksmanship success in the PR group did not increase the individual's felt acceptance (although it actually did earn him more favorable ratings from teammates), and it was not even related to his satisfaction with his own task performance. Rather, in the PR group, the individual's task satisfaction and social adjustment seemed to be closely tied to his social relationships with teammates.

In summary, individuals who perceived supportive, accepting attributes in their former teammates seemed to focus on interpersonal relationships rather than the task activity when placed into new teams. Teams composed of such persons showed good interpersonal relations but no improvement in marksmanship. In contrast, individuals in the NR group, who had not perceived former teammates as supportive, continued to be relatively unresponsive to teammates and they seemed to concentrate on their own task performance. They may, thus, have improved in marksmanship at the cost of less favorable interpersonal relations.

The two contrasting personal orientations which we have here suggested—task success and interpersonal success—are similar to concepts suggested by several other investigators. The pattern of correlations exhibited by the PR group seems fairly close to Parsons and Bales' (1953) integrative pattern, to Bales' (1958) social-emotional pattern, and to Bass' (1961) interaction orientation. The pattern characterizing the NR group, however, seems closer to Fouriezos, Hutt, and Guetzkow (1950) or Bass' self-oriented pattern than to the task orientation described by Parsons, Bales, and Bass. It was individual task success, not team task success, which seemed to be the fulcrum of individual adjustment and interpersonal relations for the NR group in this study. However, this inference must be viewed as highly speculative, since direct measures of the individual's orientation to the situation were not included in the study.

DISCUSSION

Results show that an individual's interpersonal perceptions are relatively consistent from one set of teammates to another, at least within a highly similar team situation. Apparently, as individuals enter new groups, they bring with them a relatively strong set to see others favorably or unfavorably. These perceptual tendencies, rather than the behavior of teammates, tend to determine the initial interpersonal relations in the group.

These findings, especially in conjunction with the Alexander and Drucker (1960) study, have implications for attempts to construct groups which provide a "quasi-therapeutic" environment for their members. We have shown that interpersonal perceptions which individuals bring with them into new groups are the key to initial formation of positive interpersonal relations. Alexander and Drucker found that interpersonal perceptions can be deliberately altered by experimental means, and that they ultimately aid the adjustment of teammates. At this stage of our knowledge, therefore, it would seem more profitable to attempt the assembly of "quasi-therapeutic" groups by selecting or training group members to perceive teammates in a favorable way rather than by attempting to modify their behavior toward teammates.

The present research shows that we should not expect an invariant pattern of relationships between interpersonal perceptions, adjustment, and task effectiveness. The nature of these relationships depends upon the motivation or orientation of the individual—his "stake in the game." The individual's "success" seems to be crucial to his adjustment in the group situation, but there seem to be at least two different routes to achieve this success. For some individuals, success is defined primarily in terms of effectiveness on the task: task success leads to adjustment and favorable response to teammates. For other individuals, however, personal success in a new team situation seems to be defined

in terms of the individual's social relations with his teammates.

In the initial stages of group development, members of the teammate oriented PR group did not seem to have sufficient task motivation for effective task performance. However, they tended to evaluate more favorably those teammates who did well on the task, as evidenced by the correlations between marksmanship and both esteem received and favorable behavior ratings received from teammates. Such favorable evaluation by teammates ought to be a highly desired reward for the teammate oriented person. It may, therefore, be that, over a longer period of time, teammate oriented persons would become more motivated to succeed on the task in order to gain the rewards of increased esteem from teammates.

These findings suggest that attempts to design work groups which will achieve both good member adjustment and effective task performance could best start by assessing the individual's characteristic mode of orientation to the team situation (viz., task success versus interpersonal success) as well as his characteristic ways of responding to his social environment. One type of person may well require a situation which initially provides a substantial degree of task success, while other persons may need an initially rewarding interpersonal situation. Thus, a program for development of quasi-therapeutic and effective groups appears to require a complex set of operations which attack several aspects of the problem simultaneously.

It must be emphasized that some of the findings of this study represent a posteriori hypotheses, and that the above speculations go well beyond our data. Concentrated research efforts specifically designed to test these hypotheses are indicated. Specifically, research is needed to determine the generalizability of these findings to other kinds of populations and group activities. Research is also needed to identify interaction patterns characteristic of so-called quasi-therapeutic groups. Past efforts to link interpersonal perception to interaction patterns have not been too successful. Nevertheless, this constitutes a logically compelling direction for research, if we are to amplify and extend our knowledge of the nature and effects of interpersonal relations in work groups.

SUMMARY

This study involved the experimental assembly of three-man rifle teams on the basis of prior interpersonal perceptions, and comparison of the interpersonal relations, adjustment, and effectiveness of the assembled teams subsequent to a marksmanship contest. Two types of teams were assembled: teams composed of men who had given favorable ratings to former teammates, termed the Positive Interpersonal Relations (PR) group; and teams composed of men who had not given favorable ratings to former teammates, termed the Nonpositive Interpersonal Relations (NR) group. The NR group had significantly better marksmanship scores and showed significant improvement in marksmanship while the PR group did not improve. The pattern of correlations within each of the two experimental conditions suggested that for members of the NR group success on the task was central to their situational adjustment and interpersonal relations, while for members of the PR group adjustment was related to interpersonal success rather than success on the task.

REFERENCES

ALEXANDER, S., & DRUCKER, E. H. The effects of experimentally modified interpersonal perceptions on social behavior and adjustment. Technical Report No. 9, 1960, University of Illinois, Group Effectiveness Research Laboratory, Contract DA-49-193-MD-2060, Project on interpersonal perception and the psychological adjustment of group members, Department of the Army. (Mimeo)

ALTMAN, I., & MCGRATH, J. E. A conceptual framework for the integration of small group research information. Arlington, Va.: Human Sciences Research, 1959. (HSR-TN-59/1-GN)

ALTMAN, I., & TERAUDS, ANITA. Major variables of the small group field. Arlington, Va.: Human Sciences Research, 1960. (HSR-RR-60/6-GN)

BALES, F. Task roles and social roles in problem-solving groups. In Eleanor E. Maccoby, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology*. (3rd ed.) New York: Holt, 1958. Pp. 437–447.

BASS, B. M. Comparison of behavior in groups of self-oriented, interaction-oriented, and task-oriented members. Technical Report No. 25, 1961, Louisiana State University, Contract N7 ONR 35609.

FIEDLER, F. E. *Leader attitudes and group effectiveness.* Urbana, Ill.: Univer. Illinois Press, 1958.

FIEDLER, F. E., HUTCHINS, E. B., & DODGE, JOAN S. Quasi-therapeutic relations in small college and military groups. *Psychol. Monogr.*, 1959, **73**(3, Whole No. 473).

FOURIEZOS, N., HUTT, M., & GUETZKOW, H. Measurement of self-oriented needs in discusion groups. *J. abnorm. soc. Psychol.*, 1950, **45**, 682–690.

GORDON, T. Group-centered leadership and administration. In C. R. Rogers (Ed.), *Client-centered therapy.* Boston: Houghton Mifflin, 1951. Pp. 320–383.

HUTCHINS, E. B., & FIEDLER, F. E. Task-oriented and quasi-therapeutic role functions of the leader in small military groups. *Sociometry*, 1960, **23**, 393–406.

MCGRATH, J. E. A framework for integration of small group research studies: A pilot study. Arlington, Va.: Psychological Research Associates, 1957. (PRA No. 57-20)

MYERS, A. E. The effect of team competition and success on the adjustment of group members. Unpublished doctoral dissertation, University of Illinois, 1961.

PARSONS, T. C., & BALES, R. F. The dimensions of action-space. In T. C. Parsons, R. F. Bales, & E. A. Shils (Eds.), *Working papers in the theory of action.* Glencoe, Ill.: Free Press, 1953. Pp. 63–109.

RICE, A. K. *Productivity and social organization: The Ahmedabad experiment.* London: Tavistock, 1958.

ROGERS, C. R. *Client-centered therapy.* Boston: Houghton Mifflin, 1951.

TERAUDS, ANITA, ALTMAN, I., & MCGRATH, J. E. A bibliography of small group research. Arlington, Va.: Human Sciences Research, 1960. (HSR-RR-60/2-GN)

(Received August 7, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 376-380

INDIVIDUAL DIFFERENCES, SITUATIONAL VARIABLES, AND PERSONALITY RESEARCH [1]

IRWIN G. SARASON [2]
University of Washington

Among the variables which can confound the results of human experimentation are individual differences associated with subjects and with experimenters. The question may be asked: to what extent are significant differences (or, for that matter, the lack of significant differences) between experimental treatments attributable, at least in part, to the interaction of uncontrolled individual difference variables with the manipulated variables?

Although systematic attempts at answering this question are lacking, there is increasing evidence which suggests the necessity for making such attempts. Several studies have uncovered evidence which indicates that the sex and personality characteristics of both subjects and experimenters may well influence performance in experimental situations (Binder, McConnell, & Sjoholm, 1957; Campbell, 1960; Krasner, 1958; Sarason, 1958; Taffel, 1955; Verplanck, 1955). The present research sought to go one step further by manipulating in the same experiment three individual difference variables associated with subjects and experimenters. These were sex of experimenter and hostility of subjects and experimenters as defined by scores on a paper-and-pencil test. In order to carry out the research it was necessary to have not only a sample of subjects, but also a sample of experimenters, and it was further necessary to train the experimenters in the experimental procedure.

It seemed desirable to employ this technique in a situation in which the degree of interpersonal contact between subjects and experimenter would typically be considerable. The procedure selected was one employed in studies of verbal conditioning. Following an operant period, subjects' responses falling within a previously defined response class were reinforced by the experimenter. The response class, in this instance, was verbs of a mildly hostile connotation. Because of the nature of the response class employed, it seemed particularly appropriate to incorporate the variables of hostility scores of subjects and experimenters into the experimental design.

Since the extent to which individual differences might influence performance could be affected by situational aspects of the experimenter-subject interaction, two additional variables were studied. One of these was the degree of face-to-face contact between subjects and experimenter. The other was the kind of reinforcement given the subject. Half of the subjects performed in face-to-face contact with the experimenter; the other half had no visual contact with the experimenter during the verbal conditioning series. Each of these two groups was further divided with half of each group receiving verbal reinforcement and the other half receiving nonverbal reinforcement.

There were, then, five independent variables in this research: hostility scores of experimenters, hostility scores of subjects, sex of experimenter, personal or impersonal experimenter-subject contact, and verbal or nonverbal reinforcement.[3]

The aim of the research was to determine to what extent these variables, alone and in interaction with each other, would influence the use of hostile verbs by subjects.

[1] This research was supported by a research grant (G7165) from the National Science Foundation.

[2] The writer is indebted to William Crow and James Minard for their assistance in carrying out this research.

[3] It had originally been hoped that the sex of subject variable could also be incorporated into the research design as a sixth variable. Since it did not prove to be feasible to manipulate six independent variables in the present study and since pilot studies had failed to suggest significant effects for the sex of subject variable, this variable was not studied.

METHOD

Subjects

The subjects were 160 students enrolled in introductory psychology courses at the University of Washington. All subjects were under 25 years of age. Prior to participation in the verbal conditioning experiments, they had, in a classroom situation, been administered the Autobiographical Survey (Sarason, 1958). This survey contains the Hostility scale used in the present study. The mean of the Hostility (H) scores for the subjects was 8.17 and the standard deviation was 3.20. The subjects were divided into high and low H groups, with high scorers having scores of 8 and above (mean of 10.05), and low scorers having scores of 7 and below (mean of 4.71). The group of subjects consisted of 68 males and 92 females.

Experimenters

The experimenters were 20 undergraduate students at the University of Washington. All were under 25 years of age. They had taken at least one course in psychology and had taken the Autobiographical Survey in a classroom situation. The experimenters were divided into high and low H groups on the basis of scores of 10 or above for the high H experimenters and 5 or below for low H experimenters.

Potential experimenters were contacted initially by telephone and were asked if they would like to obtain some experience in conducting psychological research. They were told that they would be paid at the rate of $1.25 per hour during their training period and during their work as experimenters. Little difficulty was encountered in obtaining the services of the 10 male and 10 female experimenters needed. Of the potential experimenters who declined to participate in the research, however, almost all had low H scores.

Training of Experimenters. Each experimenter's training was begun by his serving as a subject in a verbal conditioning session. Following this, he was informed of the nature of verbal conditioning research in general and the particular procedure used in the experiment in which he would serve as an experimenter. The experimenters were then given opportunities to practice on the research assistant conducting the training sessions and on each other. After being given as much opportunity to practice as seemed necessary, the experimenter then would "run" the research assistant as a subject. Following this, the research assistant would, if it seemed desirable, observe the experimenter through a one-way screen in an experimental session. If, at this point, an experimenter demonstrated competence in the experimental procedure, he was permitted to begin his experimental work. Each experimenter conducted eight subjects through the experimental procedure, one in each of eight experimental conditions.

Although all experimenters were fully informed as to the nature of the procedure, they were not told that the sex of experimenters variable and the variable of H scores of experimenters were being manipulated. When, at the conclusion of the experiment, the experimenters were told that these variables had been manipulated, they all indicated that they had been unaware of this manipulation.

Procedure

For all subjects there were 20 operant level and 80 conditioning trials. The response class reinforced was that of verbs with mildly hostile connotation. Each trial consisted of presentation to the subject of a $3'' \times 5''$ card on which were typed three pronouns and two verbs, one of which was neutral and the other of which was mildly hostile in connotation. Stimulus cards were stacked in a pile of 100. This pile of 100 cards was subdivided into five piles of 20 cards each. Before presenting cards to the subjects each of these piles of 20 was shuffled in order to randomize order effects. The task for the subject was to make up a sentence using one pronoun and one verb appearing on each card. The following instructions were given to each subject:

> You will see a series of cards one at a time. Each card contains three pronouns and two verbs. For each card make up a sentence using one of the pronouns at the top of the card and one of the verbs at the bottom.

For a detailed description of the procedure used in selecting the hostile and nonhostile verbs reference should be made to the paper of Binder et al. (1957).

In addition to the individual difference variables of sex of experimenter, hostility score of experimenter, and hostility score of subject, there were two experimentally manipulated variables. These were the type of reinforcement employed by the experimenter, and the nature of the experimenter-subject interaction.

The subjects were reinforced for usage of hostile verbs either verbally or nonverbally. Verbal reinforcement was administered by the experimenter saying "Good" in his normal speaking voice. Nonverbal reinforcement was administered with a 500-millisecond flashing of a 60-watt, blue light bulb located approximately 2′ from the subject's head, out of his line of vision but easily noticeable to all subjects. This light was activated by a foot pedal which the experimenter could press. All reinforcements, verbal and nonverbal, were given immediately upon completion of the subject's sentences.

The interactional variable was represented as a personal-impersonal dimension. Common to all subjects were the following conditions. The experiment was conducted in a room measuring $12' \times 12'$. A $3' \times 5'$ table separated subject and experimenter and both were seated throughout each session. The stimulus cards were presented through an aperture in a screen which shielded the experimenter's recording of the subject's responses from the subject's view. For the personal condition a screen similar to

the one employed by Binder et al. (1957) was used. This screen was only 12″ tall and permitted experimenter and subject to observe each other's facial expressions.

In the impersonal situation, the screen was 3′ tall and had supporting sides extending perpendicularly for 2′, thus effectively screening out all visual interaction between subject and experimenter. In addition, in this impersonal situation, a screen comparable in type to the one just described was used from the top of the table to the floor.

At the end of the experimental session, a series of questions was asked of each subject. These questions concerned what the subject thought the experiment was about, how he selected words for use in sentences, whether or not he noticed the reinforcement (either Good or the light flash), and if so what interpretation he made of the reinforcement. Seven of the 160 subjects were able to verbalize the contingency between reinforcement and the subject's usage of mildly hostile verbs. Comparison of these subjects with the other subjects indicated essentially similar learning curves.

RESULTS

The data for the 100 trials were subjected to a repeated measurement analysis of variance. The five variables manipulated in the study and the 100 trials, considered as five blocks of 20 trials each, were the six variables included in the analysis. Table 1 contains a summary of this analysis of variance.

Three *F*s for the uncorrelated data attained significance. The simplest of these effects was that for the verbal-nonverbal reinforcement comparison ($p < .01$). The mean number of hostile verbs used by the subjects reinforced with Good was 54.7; the mean for the subjects reinforced by a light flash 40.2. The interaction between hostility scores of experimenters and hostility scores of subjects was found to be significant at the .05 level. This

TABLE 1

ANALYSIS OF VARIANCE OF NUMBER OF MILDLY HOSTILE VERBS USED BY SUBJECTS

Source	*df*	*MS*	*F*	Source	*df*	*MS*	*F*
Sex of experimenter (A)	1	106.57	1.65	Trial blocks (T)	4	25.24	3.87***
Hostility of experimenter (B)	1	231.12	3.58*	T × A	4	6.33	<1.00
Hostility of subject (C)	1	26.64	<1.00	T × B	4	.71	<1.00
Personal versus Impersonal condition (D)	1	7.57	<1.00	T × C	4	2.60	<1.00
				T × D	4	3.04	<1.00
Verbal versus Nonverbal reinforcement (E)	1	1693.62	26.24***	T × E	4	21.39	3.28**
				T × A × B	4	5.29	<1.00
A × B	1	208.08	3.25*	T × A × C	4	12.36	1.89
A × C	1	31.21	<1.00	T × A × D	4	17.36	2.66*
A × D	1	20.48	<1.00	T × A × E	4	.33	<1.00
A × E	1	2.65	<1.00	T × B × C	4	4.17	<1.00
B × C	1	276.13	4.313**	T × B × D	4	8.72	1.34
B × D	1	.13	<1.00	T × B × E	4	8.18	1.25
B × E	1	56.23	<1.00	T × C × D	4	.11	<1.00
C × D	1	75.65	1.18	T × C × E	4	6.12	<1.00
C × E	1	98.00	1.53	T × D × E	4	3.84	<1.00
D × E	1	44.18	<1.00	T × A × B × C	4	5.09	<1.00
A × B × C	1	.29	<1.00	T × A × B × D	4	9.87	1.51
A × B × D	1	33.60	<1.00	T × A × B × E	4	5.60	<1.00
A × B × E	1	.55	<1.00	T × A × C × D	4	4.59	<1.00
A × C × D	1	81.85	1.28	T × A × C × E	4	9.37	1.43
A × C × E	1	36.12	<1.00	T × A × D × E	4	4.66	<1.00
A × D × E	1	5.44	<1.00	T × B × C × D	4	9.37	1.47
B × C × D	1	285.55	4.46**	T × B × C × E	4	4.54	<1.00
B × C × E	1	28.32	<1.00	T × B × D × E	4	7.94	1.22
B × D × E	1	1.62	<1.00	T × C × D × E	4	3.29	<1.00
C × D × E	1	172.98	2.70	T × A × B × C × D	4	.45	<1.00
A × B × C × D	1	65.87	1.02	T × A × B × C × E	4	7.64	1.17
A × B × C × E	1	16.02	<1.00	T × A × B × D × E	4	5.07	<1.00
A × B × D × E	1	136.19	2.13	T × A × C × D × E	4	7.85	1.24
A × C × D × E	1	18.31	<1.00	T × B × C × D × E	4	3.20	<1.00
B × C × D × E	1	102.77	1.60	T × A × B × C × D × E	4	2.28	<1.00
A × B × C × D × E	1	111.65	1.74	Error (w)	512	6.53	
Error (b)	128	64.02					

* $.05 < p < .10$.
** $p < .05$.
*** $p < .01$.

interaction was attributable to the differential performance of high H subjects run by high and low H experimenters. The mean number of hostile verbs used by high H subjects run by high H experimenters was 54.0; the mean for high H subjects run by low H experimenters was 42.7. The mean numbers of hostile verbs used by low H subjects run by high and low H experimenters were 46.3 and 46.8, respectively.

The triple interaction which was significant ($p > .05$) was that involving H scores for experimenters and subjects and the personal-impersonal contact variable. This interaction was a result of the use of a greater number of hostile verbs by low than by high H subjects run by low H experimenters under the impersonal condition (means of 50.8 and 37.7, respectively). Under all other conditions high H subjects used more hostile verbs than did low H subjects.

Two *F*s which approached significance ($.05 < p < .10$) were those for H scores of experimenters and the interaction of H scores of experimenters with sex of experimenters. These effects suggested that high H experimenters elicited more hostile verbs from subjects than do low H experimenters and that this was especially true for high H male experimenters.

Three of the *F*s involving correlated measures were significant. The *F* for trials was significant at the .01 level, indicating increase in usage of hostile verbs over trials. The mean number of responses within the reinforced class increased consistently over the five blocks of trials. The significant ($p < .05$) interaction between trials and the verbal-nonverbal reinforcement variable indicated differential performance over trials for verbally and nonverbally reinforced groups. Separate analyses of these two groups demonstrated a significant *F* for trials for the verbally reinforced subjects but not for the nonverbally reinforced subjects. Thus it would appear that acquisition occurred for the former but not for the latter group.

Significant at the .05 level was the triple interaction involving sex of experimenter, the personal-impersonal condition, and trials. This *F* seemed to be due to differential performance of subjects run by male experimenters during Block 2. For this block, subjects run by male experimenters used a greater number of hostile verbs under the impersonal than under the personal condition. For all other groups for all trial blocks just the opposite was the case.

DISCUSSION

The present findings demonstrate, at least for this type of experiment, that when individual difference and situational variables are systematically manipulated they have nonchance influences on performance. It would seem reasonable to draw the inference that prediction of experimental outcomes could be considerably improved when the hypotheses take into account relevant individual and situational variables as well as the independent variable being manipulated. Because of this it would seem important in future research to systematically direct attention to individual difference and situational variables which can become theoretically meaningful. One wonders to what extent, in the present psychological literature, disparate experimental findings may be due to lack of control of such variables.

The need to better integrate research in personality with experimental work is seen when the results of the present experiment are viewed in light of one performed by Binder et al. (1957). These investigators found, for the same experimental task used in the present study, that one experimenter, a female, seemed to obtain greater response change than did the other experimenter, a male. As these investigators recognized, it would be a matter of sheer speculation to hypothesize the basis of this obtained difference. The present results suggest that the difference between the Binder et al. experimenters which is relevant to their results is probably not the sexes of experimenters alone, but probably some complex of interpersonal characteristics.

The importance of the influence of the type of reinforcement on verbal responding was demonstrated in the present experiment. No indication of a conditioning-like effect occurred for the subjects reinforced with the light flash. A significant increase in usage of hostile verbs was observed for the

verbally reinforced subjects. The inability of the light flash to function as a positive reinforcer for a sentence construction task is consistent with previous findings of Taffel (1955) and Nuthmann (1957). On the other hand, for other tasks, Greenspoon (1954) and Sidowski (1954) have found nonverbal reinforcers to be effective. The problem of what constitutes a reinforcing event for different types of response classes would seem to merit further research.

Although H scores did not interact with trials, there was evidence that they did relate to the subject's usage of hostile verbs. Especially interesting was the interaction of the subject's and the experimenter's H scores. The highest mean number of hostile verbs was obtained from the high H subjects seen by high H experimenters. One possibility suggested by these results is that the level of the subject's responsiveness may be influenced by the degree of experimenter-subject compatibility or similarity. If this were true, it would be important to determine in what kinds of situations this is a factor affecting performance.

SUMMARY

The effects of five variables on the subject's verbal responses were studied in the present research. The five variables were: sex of experimenters, the experimenter's scores on a hostility scale, the subject's scores on the same hostility scale, verbal versus nonverbal reinforcement, and face-to-face versus impersonal experimenter-subject interaction. The 160 subjects were shown 100 3″ × 5″ cards on each of which were typed three pronouns and two verbs, one of which was mildly hostile in connotation, the other of which was neutral. The subject's task was to construct sentences using one of the pronouns and one of the verbs. Reinforcement, either the experimenter's uttering "Good" or a light flash, followed the subject's usage of mildly hostile verbs in sentences.

The results showed that for all subjects reinforcement did lead to increases in the subject's usage of hostile verbs. However, when the verbally and the nonverbally reinforced subjects were compared, the former but not the latter subjects showed an acquisition effect. Hostility scores of subjects and experimenters significantly interacted with each other with the highest number of hostile responses coming from high hostile subjects seen by high hostile experimenters. Other, more complex, interactions were also obtained.

The findings of the experiment were interpreted as demonstrating the important roles of individual difference and situational variables in psychological research.

REFERENCES

BINDER, A., MCCONNELL, D., & SJOHOLM, NANCY A. Verbal conditioning as a function of experimenter characteristics. *J. abnorm. soc. Psychol.*, 1957, **55**, 309–314.

CAMPBELL, J. M. Verbal conditioning as a function of the personality characteristics of experimenters and subjects. Unpublished doctoral dissertation, University of Washington, 1960.

GREENSPOON, J. The effect of two nonverbal stimuli on the frequency of members of two verbal response classes. *Amer. Psychologist*, 1954, **9**, 384. (Abstract)

KRASNER, L. Studies of the conditioning of verbal behavior. *Psychol. Bull.*, 1958, **55**, 148–170.

NUTHMANN, ANNE M. Conditioning of a response class on a personality test. *J. abnorm. soc. Psychol.*, 1957, **54**, 19–23.

SARASON, I. G. Interrelations among individual difference variables, behavior in psychotherapy, and verbal conditioning. *J. abnorm. soc. Psychol.*, 1958, **56**, 339–344.

SIDOWSKI, J. B. Influence of awareness of reinforcement on verbal conditioning. *J. exp. Psychol.*, 1954, **48**, 355–360.

TAFFEL, C. Anxiety and the conditioning of verbal behavior. *J. abnorm. soc. Psychol.*, 1955, **51**, 496–501.

VERPLANCK, W. S. The control of the content of conversation: Reinforcement of statements of opinion. *J. abnorm. soc. Psychol.*, 1955, **51**, 668–676.

(Received August 8, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 381–386

EFFECTS OF AMBIGUITY OF ENVIRONMENTAL CUES UPON LEARNING TO IMITATE

JOHN W. McDAVID [1]

University of Miami

The Miller and Dollard (1941) analysis of social learning and imitation within the framework of reinforcement theory has engendered extensive theoretical analysis and empirical investigation of the imitative learning process (e.g., Church, 1957; Kanareff & Lanzetta, 1958; McDavid, 1959; Rosenblith, 1961; Wilson, 1958). However, critics of this position have argued that the strict reinforcement analysis outlined by Miller and Dollard tends to ignore the importance of information and conceptual solutions of the problem or task in which imitation occurs. This issue is obviously a crucial one meriting considerably more direct experimental attention than it has received. Asch's (1952) criticisms notwithstanding, the Miller and Dollard analysis does take indirect note of the informational utility of imitation in social learning. Imitation and learning by observation, they suggest, may direct the imitator's attention toward relevant environmental cues which he otherwise might ignore or be unable to interpret.

In order to explore further the role of information in the imitation learning situation, a modification of the original Miller and Dollard (1941) experimental paradigm was devised to investigate the extent to which the to-be-imitated response may tend to direct the imitator's attention toward or away from various discriminable cues in the physical environment, and vice versa. The subjects employed in the present study were 4-year-old preschool children, and the experimental leaders were young adults rather than peers. The reward to be found was a marble token to be exchanged later for a "prize" and the discrimination problem itself was somewhat more complex than that employed by Miller and Dollard, in that it afforded three, rather than two, response alternatives distinguished by salient environmental cues (color and position). Experimental procedures were arranged to vary the information about the potential relevance of color cues conveyed by the observed choices of the leader, while holding constant the administration of reward to the subject only for direct imitation of the leader's choice behavior. The leader was pre-instructed to make choices arranged to be either consistently (100%), partially (67%), or randomly (33%) associated with the environmental (nonsocial) color discrimination.

A study of "nonsocial" discrimination in a concept formation task by Gormezano and Grant (1958) provides an analysis which may be extrapolated to social situations in which the revelant cue for concept formation or problem solution is the response of another person. This investigation explored the effects of varying reinforcement of irrelevant dimensions in a card sorting concept formation task, and demonstrated that as the ratio of reinforcement for sorts on an irrelevant dimension increased, with consistent reinforcement for sorts on a single relevant dimension (100%), difficulty in attainment of the appropriate conceptual solution increased. These authors interpret their findings in support of a "mediational" or "selective attention" interpretation of conceptual learning—an interpretation which is congruent with the implications of the Miller and Dollard (1941) discussion of the role of imitation in redirecting attention toward relevant environmental cues. According to a similar analysis of the learning situation in the present study, then, one would expect greater difficulty in learning to imitate under the condition of 67% association than under

[1] The data reported here were collected while the author was a member of the faculty of the Iowa Child Welfare Research Station, State University of Iowa.

Jerry Alpern, Lois Hendricks, and Ray Ditrich were generous in their assistance as experimental leaders.

the condition of 33% association between color and leader's choice response. Learning should proceed most rapidly under conditions of complete (100%) association of color and leader's behavior, both in this case being relevant cues.

METHOD

Subjects

Subjects for this study included 30 children between the ages of 47 and 63 months, with a median age of 56 months. There were 15 males and 15 females, and all were enrolled in either the Iowa Child Welfare Research Station laboratory preschool or the Iowa City Parents' Cooperative Preschool. This population represents predominantly middle-class families who are relatively permanent residents of the community. The children were well accustomed to participating in experimental studies executed in the preschools.

The experimental leaders were young adult graduate students with whom the experimental subjects were familiar by sight. A male adult served as leader for male subjects, and a female adult served as leader for female subjects.

Procedure

Apparatus. The apparatus employed in this study was somewhat more complex than that of Miller and Dollard (1941). The problem box consisted of a black wooden box approximately 10 inches deep and 10 inches × 18 inches on the front face, placed on a low table at chest height to the seated subject. The face of the box contained three 2 inch × 2 inch opal glass apertures spaced 4 inches apart, with a small momentary contact response button centered 2 inches below each aperture. The apparatus was constructed to dispense a marble automatically following each correct button-pressing response into a small wire trough at the bottom center of the box. Red, yellow, and green electric bulbs behind each opal glass aperture afforded presentation of any of these colors in any of the three apertures. All experimental controls for the problem box were contained in a central remote control panel in an adjacent observation booth arranged so that the experimenter could observe the experimental procedure without himself being physically present in the room with the subject and the leader.

Task and Instructions. The subject's task was to attempt to obtain as many marbles as possible by pressing one of the three buttons under the color apertures on each trial. The experimental series included 24 trials. On each trial, the adult leader made his response in the subject's presence, and the subject was instructed to make his response immediately following. There was approximately a 10-second interval between trials, during which time the experimenter reset the presentation of color cues and designation of correct response for automatic dispensation of the marble reward. To introduce the subject to the experimental session, the experimenter took him from the preschool playroom, went to the leader (standing unoccupied in the playroom) and asked for his assistance in "playing a game," and then led both the subject and the leader to the experimental room. Before leaving the room to enter the observation-control booth, the experimenter read the following instructions to both the leader and the subject:

> Subject [subject's name], and Leader [leader's name], this is a game for two people to play at the same time. You will each take turns playing the game, and if you play it the right way, you will win some marbles. When we are through, if you have enough marbles, you can trade them for one of these toys. Both of you can win toys, so let's pick out the ones you would like to have.
>
> Now I will show you how to play the game to get marbles. Sit on these chairs right here: Subject, you sit on this one [nearer center of apparatus]; Leader, you sit on that one [off-center to the left]. You see, this box has three windows, and each one is a different color. Under each window there is a little button. Every time the lights come on, each of you will get a turn to push one of the buttons. If you push the right button, a marble will come out down here. Now there is only one right button to push, and you must try to push the correct button every time so that you can get as many marbles as possible. Here is a box for each of you to keep your marbles in until the game is over.
>
> Leader, suppose you go first each time. And Subject, you can go second. Now remember, Leader, you must wait until the lights come on and then take your turn. Push only one button, and push it just one time. Then, Subject, when Leader gets through, you will take your turn. When both of you have had a turn, the lights will go out and you will wait until the lights come on again for a whole new turn of the game. Do you understand how to play the game now?
>
> We will make a rule not to talk to each other about the game until it is over. Now I will go into the other room and start the game. But when the game is over, I will come back and we will see if you have enough marbles to win the toy you picked. Are we ready to start the game now?

When the subject indicated understanding of the experimental task, the experimenter went into the adjacent booth and began the training series, which continued without interruption through 24 trials. Presentation of the six possible color-position arrangements was randomized within the series, so that each of the possible combinations appeared once in each block of six trials. While the 24-trial series was not sufficiently long for all subjects to solve the problem, this length was chosen for optimum sustained interest and investment on the

subject's part, and it is adequate to reflect differential effects in learning to make the consistently reinforced imitative response.

Conditional Manipulations. Five male and five female subjects were trained under each of three conditions which varied in the degree of association between the leader's response and the salient differential color cues associated with each response. Reinforced position was varied randomly for all subjects within all aspects of each of the three conditions. In Condition 100, the leader consistently responded to the yellow color cue, so that the subject was reinforced simultaneously for imitation of the leader and for approach response to the yellow cut. In Condition 67, the leader responded four times in every block of six trials to the yellow cue, once to the red cue, and once to the green cue; hence, the subject was reinforced consistently for imitation of the leader, but 67% of the time for responding to the yellow cue, 17% of the time for responding to the green cue, or 17% of the time for responding to the red cue. In Condition 33, the leader responded twice within each block of six trials to each of the three color cues; hence, the subject was randomly and nondifferentially reinforced for color based responses, while being consistently reinforced for imitation of the leader. The leader received reinforcement for each of his responses in all three conditions.

RESULTS AND DISCUSSION

Scores based on the number of reinforced (imitative) responses made by the subjects in each block of six trials were recorded and submitted to a mixed design analysis of variance, in which the subject's sex and the experimental condition (100%, 67%, or 33% color imitation association) constituted between-subjects factors, and trial blocks constituted a within-subjects factor. Figure 1 summarizes these scores graphically, and the results of the analysis of variance are summarized in Table 1.

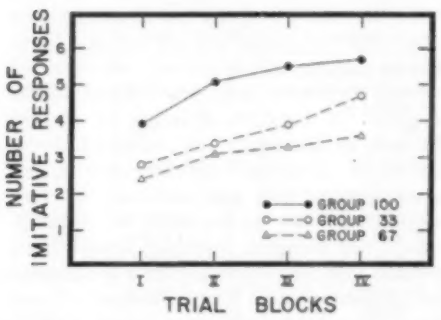

FIG. 1. Number of imitative responses by trial blocks for three experimental conditions.

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE OF IMITATIVE RESPONSES DURING TRAINING SEQUENCE

Source	*df*	*MS*	*F*
Between subjects	29		
Experimental condition (C)	2	39.900	6.609*
Sex (S)	1	1.200	—[a]
C × S	2	0.400	—
Error (b)	24	6.038	
Within subjects	90		
Trial blocks (T)	3	14.722	14.933**
T × C	6	0.556	—
T × S	3	0.289	—
T × C × S	6	3.656	3.723*
Error (w)	72	0.982	
Total	119		

[a] *F* ratio less than unity.
* $p < .01$.
** $p < .001$.

The sex factor was not significant, but the experimental conditions factor was significant ($p < .01$) as expected. The condition (Group 100) in which the leader's behavior was consistently associated with a particular color cue clearly afforded the greatest facility for learning to imitate the leader, and the mean number of correct responses for this group (20.2) was significantly greater than that for Group 67 (12.4) as well as that for Group 33 (14.8) at the .01 confidence level ($t = 14.2$ and 9.8, respectively). The mean total correct responses for Group 67 (12.4) was significantly less than that for Group 33 (14.8), also at the .01 level of confidence ($t = 4.4$). These findings clearly support the expectation that Group 100 would learn the problem most easily, with Group 67 having even more difficulty than Group 33 in learning to imitate.

The trials factor was significant ($p < .001$), indicating that learning did take place during the training sequence. Neither of its interactions with the between-subjects factors were statistically significant, and the similarity of the forms of the learning curves for the three experimental conditions is apparent in Figure 1. The means for all groups combined for the four trial blocks were 3.03, 3.87, 4.20, and 4.70, in consecutive order, and all of the differences between individual

pairs of trial block means were significant ($p < .01$) with the exception of the differences between Blocks 2 and 3 and Blocks 3 and 4.

The triple interaction (Trials × Experimental Conditions × Sex) was unexpectedly significant ($p < .01$). This significance appears to be due essentially to differences in the forms of learning curves over trials for males and females in Group 67. The females made significantly fewer ($p < .02$) correct responses during the first block of six trials ($M = 1.6$) than did males ($M = 3.2$) in this group, while the females made significantly more ($p < .02$) correct responses during the last block of six trials ($M = 4.4$) than did the males ($M = 3.0$). The males in this group show virtually no learning at all over the training series (*M*s for the four consecutive trial blocks were 3.2, 3.2, 2.6, and 3.0). The females, however, show evidence of learning comparable to the other experimental groups (*M*s for the four consecutive trial blocks were 1.6, 3.0, 3.8, and 4.4). While this interaction effect has some particularly interesting implications, it should not be overinterpreted without replication. These data are based upon only five subjects of each sex in each condition. Nevertheless, the fact that male subjects show evidence of learning to imitate during the 24-trial training sequence when physical cues are either consistently or randomly associated with imitation, but not when physical cues are partially systematically associated with imitation, suggests that boys may be disposed to attend to such physical cues as color in preference to cues derived from the behavior of other people (social cues). It may be that girls are relatively more disposed to attend to such social cues for directives for behavior, thus suffering less interference from the ambiguously irrelevant color cue in Condition 67. This interpretation is consistent with an interpretation of borderline sex differences in imitation in an earlier study (McDavid, 1959) which follows from the suggestion of Terman and Tyler (1954) that young girls tend to be more "person oriented" and boys more "thing oriented."

In order to further explore differences in imitative behavior between Group 67 and Group 33, an additional analysis was undertaken. It should be noted that on two-thirds of the trials for Group 67, a "correct" response may be based on *either* imitation *or* color. Thus, a subject who might have accepted the partially correct color solution as a consistent mode of behavior would necessarily be recorded here as having imitated two-thirds of the time. As a result, those trials on which the leader did not respond to the partially correct color cue were designated *test trials*, and imitative responses on these trials were recorded and converted into percentages. These percentage scores were recorded for the first and last halves of the training series separately, and were compared to percentage scores for imitation for Group 33. (In the latter case, no single color cue was reinforced any more often than any other, so that the percentage scores for imitation for this group were derived from all trials.) These percentage scores were submitted to a mixed design analysis of variance, with subject's sex and experimental condition (Groups 67 and 33) constituting between-subjects factors and trial blocks (first half versus last half) constituting a within-subjects factor. The results of this analysis are summarized in Table 2. Only the effect of experimental conditions was significant ($p < .001$). While there is some heterogeneity of variance apparent, Bartlett's test was negative, and the significant *F* ratio

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE OF PERCENTAGE TEST TRIAL IMITATION FOR GROUP 67 AND GROUP 33

Source	*df*	*MS*	*F*
Between subjects	19		
Experimental condition (C)	1	19,536.4	21.556**
Sex (S)	1	8.3	—[a]
C × S	1	34.8	—
Error (b)	16	906.3	
Within subjects	20		
Trial blocks (T)	1	1,000.0	4.119
T × C	1	1,000.0	4.119
T × S	1	108.7	—
T × C × S	1	762.2	3.139
Error (w)	16	242.8	
Total	39		

[a] *F* ratio less than unity.
** $p < .001$.

falls beyond an extreme level of confidence. The mean percentage of imitative responses on test trials for Group 67 was 17.5%, as compared to 61.7% for Group 33. This sharp difference suggests further that many subjects in Group 67 may have accepted the partially correct color solution to the discrimination problem as a consistent mode of behavior. The proportion of imitative responses on test trials for this group (17.5%) is considerably below chance level occurrence, as well as below that predicted from a strict stimulus compounding analysis of the learning situation. While the trials effect, as well as the Trials × Experimental Conditions interaction, fell short of statistical significance ($.10 > p > .05$), there is an evident increase in imitation during the training series for Group 33 (from 51.7% to 71.7%, by halves), with no increase at all for Group 67 (17.5% and 17.5%).

While the 24-trial training series did not allow all subjects to learn the imitative response to a criterion of six consecutive correct responses, the three experimental conditions differ sharply in terms of the proportion of subjects within each group who attained this criterion. The proportion of subjects meeting this criterion in Group 100 was 90% (five males and four females); in Group 33 it was 50% (two males and three females); and in Group 67 it was 10% (one female). These differences further emphasize the differential effects of the training conditions on imitative behavior.

The relative levels of imitative response observed under each of three experimental training conditions employed here offer evidence that partial but incomplete association between the leader's behavior and specific environmental cues tends to reduce the frequency of imitative response on trials in which imitation and nonsocial or environmental cues are not associated. It appears that this condition may lead to selective redirection of attention to the related color cues, with a tendency to ignore the leader's behavior as a unique single cue. This observation is consistent with the Gormezano and Grant (1958) assumption of "mediational" processes of selective attention in conceptual learning.

Since partial but incomplete association between an imitated leader's behavior and environmental cues typifies the everyday occurrence of imitative or observational learning in child rearing and social influence, this paradigm may be typical of commonly observed processes of imitation outside the laboratory.

SUMMARY

In order to explore certain aspects of the role of the informational content of another person's observed behavior upon the process of learning to imitate that person, an extension of the Miller and Dollard (1941) imitation paradigm was formulated. Preschool subjects were trained to imitate an adult leader in a three-choice discrimination learning problem in which imitation was consistently rewarded with a marble token, while salient color cues were associated with the leader's response at three levels of ambiguity. For one group of subjects, the leader consistently responded positively to the yellow cue, resulting in 100% coincidence of color and imitative (social) cues; for a second group, the leader responded positively two-thirds of the time to the yellow cue, and one-sixth of the time to each of two other color cues, resulting in 67% coincidence of color and imitative cues; for a third group, the leader responded randomly to each of the three color cues, resulting in 33% coincidence of color and imitative cues.

During a 24-trial training sequence, subjects evidenced an increase in imitative responses. However, the degree of association between the relevant (imitative) and a salient irrelevant (color) cue affected levels of performance significantly, such that learning of the imitative solution to the problem took place most readily when color and imitation were consistently associated, with somewhat more difficulty when color and imitation were randomly associated, and with even greater difficulty when color was partially associated with imitation two-thirds of the time.

The results are construed as evidence that direction of the imitator's attention to cues in the physical environment is an important aspect of the imitation process, and that the

degree of ambiguity of such environmental cues has a direct effect on the learning process. Of several plausible theoretical analyses of the phenomena, it is felt that a "mediational" or "selective attention" interpretation of conceptual learning is most consistent with the empirical observations.

REFERENCES

ASCH, S. E. *Social psychology*. New York: Prentice-Hall, 1952.

CHURCH, R. M. Transmission of learned behavior between rats. *J. abnorm. soc. Psychol.*, 1957, **54**, 163–165.

GORMEZANO, I., & GRANT, D. A. Progressive ambiguity in the attainment of concepts on the Wisconsin Card Sorting Task. *J. exp. Psychol.*, 1958, **55**, 621–627.

KANAREFF, VERA T., & LANZETTA, J. T. The acquisition of imitative and opposition responses under two conditions of instruction induced set. *J. exp. Psychol.*, 1958, **56**, 516–528.

MCDAVID, J. W. Imitative behavior in preschool children. *Psychol. Monogr.*, 1959, **73**(16, Whole No. 486).

MILLER, N. E., & DOLLARD, J. *Social learning and imitation*. New Haven: Yale Univer. Press, 1941.

ROSENBLITH, JUDITH F. Imitative color choices in kindergarten children. *Child Develpm.*, 1961, **32**, 211–223.

TERMAN, L. M., & TYLER, L. E. Psychological sex differences. In L. Carmichael (Ed.), *Manual of child psychology*. (2nd ed.) New York: Wiley, 1954. Ch. 17.

WILSON, W. C. Imitation and the learning of incidental cues by preschool children. *Child Develpm.*, 1958, **29**, 393–398.

(Received September 1, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 387–394

THE PERFORMANCE OF PROCESS AND REACTIVE SCHIZOPHRENICS AND BRAIN INJURED SUBJECTS ON A CONCEPTUAL TASK

THOMAS A. TUTKO
Northwestern University

AND

JANET TAYLOR SPENCE
Veterans Administration Hospital, Iowa City

From the inception of investigations into the group of patients called "schizophrenics," distinctions have often been made between those individuals in whom the onset of the psychosis has been early but slow and insidious, with a correlated history of poor social adjustment, as opposed to those individuals for whom the premorbid history has been relatively adequate, onset of the psychosis more rapid and frequently precipitated by some stressful event. Those with a slow onset, it has often been noted, show somewhat different symptomatology (e.g., dull, inappropriate affect) and a poorer prognosis than those with a more precipitous onset and a superior premorbid social history. A variety of terms has been employed to distinguish between these two extreme groups: dementia praecox versus schizophrenia, typical versus atypical, schizophrenic versus schizophreniform, etc. The labels currently most in favor (and which will be employed here) refer to Process versus Reactive schizophrenics, the Process schizophrenic being the one in whom the premorbid adjustment is poor and onset slow.

It has frequently been suggested that different etiological factors may be responsible for these two "types" of schizophrenia. One specific proposal that has been offered is that psychogenic factors may be primarily responsible for the development of Reactive schizophrenia, while somatogenic or "organic" factors may be crucial in the development of Process schizophrenia (e.g., Bellak, 1958). Some formal support for such a suggestion may be found in a study by Brackbill and Fine (1956). These investigators scored Rorschach protocols for a number of Piotrowski's signs of organicity; no significant differences were found between a group of Process schizophrenics and a group of brain injured patients although both of these groups exhibited significantly more signs of organicity than Reactives. Negative evidence, however, has been reported by McDonough (1960) who was unable to find the similarities in performance between Process and brain injured groups on the Archimedes Spiral and Critical Flicker Fusion tasks that would be anticipated by the organicity hypothesis.

In the present study the performances of brain injured patients and Process and Reactive schizophrenics were compared on a conceptual task, namely, the Goldstein-Scheerer (1941) Sorting test. As in the investigation by Brackbill and Fine (1956), an attempt was made to demonstrate that the performance of brain injured subjects and Process schizophrenics is more similar than that of Process and Reactive subjects. More specifically, predictions were made about the type of errors that would tend to be made by Reactive schizophrenics and the type that would tend to be made by both Process schizophrenics and the brain injured.[1]

In presenting the rationale from which the scoring system used in the present study was derived, it is important to consider first the characteristics of a "correct" verbal response, as that term is typically used in investigations of the performance of schizophrenic and/or brain injured patients on conceptual tasks (e.g., Rapaport, 1945). Given a set of objects about which he is asked to form a verbal "concept," the subject is required to specify at least one property that all of the objects have in common. Obviously, for the

[1] It is possible to analyze the characteristics of adequate responses as well as inadequate, e.g., to assign them to such categories as abstract-conceptual, functional, and concrete. Although the test protocols were scored in this manner in the present study, no predictions were developed concerning the relative frequency with which the various types of adequate responses would occur in the experimental groups.

response to be considered correct the property that the subject specifies must not only be appropriate but apply to *all* of the objects rather than only to some of them. But, to be considered correct or adequate, the subject's response must not merely refer to some common property. In addition, it is typically required that the subject refer to one of the relatively *specific* and objectively obvious characteristics shared by the particular set of objects. For example, a group of metal objects may also have in common the property of being manufactured, useful to man, made of molecules, etc. The latter types of description, although they are correct, go far beyond the "lowest common denominator" of shared characteristics among the particular group of objects. Such responses (which seem to demonstrate what Cameron, 1938, 1939, calls "overinclusiveness") are typically considered to be too broad and general, and are therefore scored "incorrect" or inadequate. Second, the property specified by the subject must be based on objective characteristics of the objects, i.e., those apparent to other observers and not idiosyncratic for a particular subject. Thus, such responses as "They are all frightening" or "All are necessary for healthful understanding," are considered to be incorrect or inadequate. In the terminology employed by McGaughran and Moran (1956), the concepts must be "public" rather than "private."

Proceeding from these general statements, it is possible to divide incorrect or inadequate responses into two major categories. The first might be described as too general or nonspecific: subject is able to report similarities among the stimuli but his responses in some manner go far beyond the particular stimuli and their observable properties. The second category is in a sense the opposite of the first: subject is unable to specify any common characteristic (the objects appear unique) or specifies a property that is shared by only some of the objects. It is tempting to label the first type "abstract errors" and the second "concrete errors." Certainly the latter seem to refer to the kind of behavior that led Goldstein and his co-workers (Bolles & Goldstein, 1938; Goldstein & Scheerer, 1941) to characterize the thinking of brain injured patients as "concrete" and "stimulus bound." Among schizophrenics, it is not uncommon to find both types of inadequate responses, those that are too bound to the uniqueness of the stimuli and those that are too general and vague. Both, unfortunately, have often been labeled "concrete" so that the term becomes approximately equivalent to "poor" or "unsatisfactory."

In the present study, an attempt was made to relate the type of error made on a conceptual task to independently defined characteristics of the schizophrenic patient. The errors of Reactives, it was predicted, would predominantly be of the type that are inappropriately broad or idiosyncratic (hereafter called Expansive errors). The errors of Process schizophrenics, in contrast, were expected to reflect their inability to suggest similarities among sets of objects (hereafter called Restrictive errors). It was further predicted that these Restrictive errors would also characterize the performance of brain injured subjects so that, in this respect, the behavior of the Process schizophrenics would be more similar to this group than to Reactives.

The selection of Process and Reactive schizophrenics was based on the procedure employed by Rodnick and Garmezy and their students (1957). In their studies these investigators have rated subjects on that portion of the Phillips (1953) Prognostic Rating Scale that evaluates adequacy of premorbid adjustment.[2] As in the original vali-

[2] In addition to premorbid adjustment, the Phillips (1953) scale also contains sets of items referring to the presence of precipitating events and signs of the disorder. Data concerning the former are often difficult to obtain or assess from case history material. Further, Phillips was unable to demonstrate the prognostic significance of these items. For these reasons, subjects were not rated on this part of the scale in the present study although the presence or absence of precipitating factors is typically mentioned as a distinction between Process and Reactive schizophrenics. That part of the Phillips scale dealing with signs of the disorder, however, has been found to be prognostically significant. Since signs of the disorder have also been found to be highly correlated with premorbid adjustment (Phillips, 1953; Rodnick & Garmezy, 1957), inclusion of both of these areas for selective purposes may be unnecessarily redundant. More important for research, it is important *not* to *select* subjects so that they

dating study (Phillips, 1953), Rodnick and Garmezy have demonstrated that the Premorbid Adjustment scale has prognostic value. Schizophrenics dichotomized on this basis have also been found to differ in their performance on a number of laboratory tasks (Harris, 1957; Rodnick & Garmezy, 1957; Zahn, 1959).

METHOD

Subjects. Four groups of 15 subjects each were tested, all subjects being hospitalized male veterans drawn from Veterans Administration Hospitals at Downey, Hines, and West Side (Chicago), Illinois. The normal control subjects were ambulatory patients on TB wards; none had a history of psychiatric or neurological disorders as determined from case records. The brain injured group were drawn from patients on neurological wards; all had been given a positive diagnosis of cortical damage and none had a history of psychiatric disturbance. Members of the two schizophrenic groups were labeled as Reactive or Process according to the criteria outlined below. None of these subjects evidenced any history of neurological disorder or of alcoholism. For all groups an attempt was made to select subjects whose current admission to the hospital had been relatively recent; none of the members of the normal and schizophrenic groups had been hospitalized for more than 12 months and none of the brain injured for more than 18 months. The four groups were equated in age and education, the mean age for the total group being 35.4 years and mean number of grades completed being 11.0.

The subjects in the two schizophrenic groups were selected on the basis of ratings of adequacy of premorbid adjustment made from case history materials. The ratings were based on Area I of the Phillips (1953) Prognostic Rating Scale in which five aspects of premorbid adjustment are evaluated: recent sexual adjustment, social aspects of sexual life during adolescence and immediately beyond, recent premorbid personal relations, and history of personal relations. Cutoff scores of 18 or above (out of a possible 30) were established for members of the Process group and 14 or below for members of the Reactive group. Records of patients who had been judged to be sufficiently intact to perform the experimental task continued to be examined until the desired number of individuals whose scores fell within the specified ranges and who met the other criteria of selection were obtained. Ratings on the Phillips scale were made independently by two clinical psychologists; in all cases they agreed within the score limits set for the two schizophrenic groups.

All subjects were interviewed briefly several days before being tested to determine their willingness to serve in the study, their capacity to understand instructions and, in the case of the brain injured subjects, their ability to communicate verbally. Of the subjects who were approached, only three (all normal subjects) refused to participate in the study. Two additional subjects, both brain injured, were judged to be unable to perform the experimental task and were thus not tested.

Experimental task and procedure. The Goldstein-Scheerer Sorting test, consisting of 33 common objects, was given to all subjects according to the procedure described by Rapaport (1945). In the first part of the test procedure, the subject was presented with a series of single objects and, for each one, asked to select all of the other objects that "belong with it" and then to state the basis for his choice. In the second part of the test, the experimenter placed standardized groups of items before the subject, in each case instructing the subject "Now tell why all these belong together." There were 12 such groups, consisting of 5–12 objects or pairs of objects per group. Each group permitted a conceptual definition, e.g., all red, all metal, tools, etc. Because of the differences within and between subjects in the number of objects chosen on the first part of the test, a quantitive scoring system was difficult to devise. For this reason, only the verbal responses obtained on the second part were scored and subjected to statistical analysis in the present study.

The subject's response to each of the 12 items on Part II was first scored as being Adequate or Inadequate. A response was considered Adequate if the subject correctly pointed out some common and fairly specific attribute of the objects in the course of his answer, i.e., even if his answer also contained additional, inadequate responses. Following Rapaport's (1945) definitions and examples, the Adequate responses were also scored as being Abstract, Functional, or Concrete.

Inadequate responses were assigned to one of six error categories. The labels given to each of them and a brief description are as follows:

1. No Response. Subject is unable to offer a basis for the sort.

2. Split-narrow. Subject defines, not the total group of objects, but one or more subgroups (e.g., "these are paper and these are rubber").

3. Chaining and Object naming. Subject merely names or describes objects in the groups in a loose, nonintegrated manner.

4. Concrete Fabulization. Subject mentions objects in the group, making them a part of a story or narrative (e.g., "The man lights the cigarette with the match and then he writes on a piece of paper").

5. Abstract Fabulization. As in Concrete Fabulization, the subject's response is essentially a narrative. However, the subject does not mention specifically

---

are already different in the very respects that one is testing in a particular investigation. To avoid this problem, it seems desirable to follow Rodnick and Garmezy (who prefer to call their groups, more neutrally, Goods and Poors), in selecting Process and Reactive subjects solely on the basis of premorbid adjustment.

or goes far beyond the objects in the group (e.g., "You handle a man, give him a cigarette and check overt conduct obvious behavior").

6. Syncretistic and Symbolic responses. The subject's response specifies a common attribute among the objects that is vague and overgeneralized (e.g., for tools, "They are all useful."), or contains loose associations to the objects (e.g., for real and imitation smoking materials, "man and boy").

These categories were taken from Rapaport (1945) and with the exception of the two types of fabulization, the subject's responses were scored in accordance with his definitions of them. Rapaport includes fabulatory responses among his categories, but does not make the distinction between Concrete and Abstract narratives. For purposes of the present study, responses falling in the first four categories were designated as Restrictive errors and those in the last two (Abstract Fabulization, Syncretistic and Symbolic) as Expansive errors.

The subject's responses occasionally contained several independent bases for the sort (e.g., "All manufactured; all made of molecules"). In order to keep the total number of responses constant for each subject, each response was assigned to only one score category. As mentioned previously, if one of the bases could be considered correct, the total response was placed in the Adequate category. For responses that were inadequate, the total response was placed in only one of the error categories. Since these multiple responses were usually of the same specific type (e.g., all Syncretistic) and always of the same general type (all Restrictive or all Expansive) they constituted only a minor scoring problem.

All of the test protocols were scored by one of the experimenters (TAT) without knowledge of the group to which the subject belonged. To check the reliability of the scoring, 30 responses initially scored as Adequate and 30 initially scored as Inadequate that the experimenter considered the most difficult to assign to scoring categories were first selected. To these were added an equal number of responses from the two major categories selected essentially at random but in such a manner that an equal number of responses to each of the 12 sorts were represented among the 120 responses. These were given to two clinical psychologists, along with definitions of the various categories and a description of the particular objects, with the instructions to rate them as Adequate or Inadequate, and within each of these to assign them to subcategory. Interjudge agreement among the three raters averaged 89% for assignment to Adequate versus Inadequate, 92% for assignment of Adequate responses to Abstract versus Functional and Concrete, and 84% for assignment of Inadequate responses to one of the six error categories. When percentages of agreement as to major error category (Restrictive or Expansive) was computed, the value rose to 96%. Considering the method by which the responses to be checked were selected, the reliability of the original scoring was considered to be satisfactory.

RESULTS

The mean scores of the four groups for various scoring categories are shown in Table 1. Considering first the overall performance of the groups, it may be seen that the normal control group gave the greatest mean number of adequate responses (8.33 out of a possible 12) and the brain injured the least. Of greater interest to the present study, the Reactive and Process schizophrenics were quite similar in performance. Analysis of these data by means of t tests indicated that the performance of the normal subjects was significantly superior to each of the other groups ($p \leqq .05$). The difference between the brain injured and each of the schizophrenic groups approached significance ($p < .10 > .05$).

Each Adequate response, it will be recalled, was scored as being Abstract, Functional, or Concrete. Since the groups were not equal in total number of Adequate responses, it was not possible to compare the groups directly by examining the number of responses in each category. Instead, the percentage of the total number of Adequate responses that were scored as Abstract was determined for each subject and from these, the mean percentage for each group. This procedure, in effect, lumps the Functional and Concrete responses into a single category; since the latter occurred very rarely, it was felt that no important differences among the groups were obscured.

As may be seen in Table 1, the mean percentage values were all greater than 50, indicating that Abstract responses were given more frequently than Functional and Concrete in all groups. However, there were marked differences among the groups. As in the comparison of total number of Adequate responses, the performance of the normals was "best" (the greatest percentage of Abstract responses), and the brain injured worst. An analysis of variance indicated that the means of the groups differed significantly ($F = 3.78$, $p = .05$). However, t tests indicated that while the normals were significantly superior to each of the other groups ($p \leqq .05$), none of the other intergroup comparisons approached significance.

TABLE 1

MEANS OF VARIOUS RESPONSE MEASURES FOR THE FOUR GROUPS

Response	Group							
	Normal		Reactive		Process		Brain injured	
	M	σ	*M*	σ	*M*	σ	*M*	σ
Number Adequate	8.33	2.40	5.60	2.94	5.00	2.85	3.73	2.24
Percentage Abstract	91.93	13.41	61.27	32.55	73.40	31.95	55.40	39.95
Number Expansive errors	2.07	1.61	4.67	3.48	.87	1.02	.47	.81
Number Restrictive errors	1.60	1.74	1.67	2.05	6.13	3.18	7.87	2.41
Percentage Restrictive errors	36.87	36.14	26.07	30.21	80.73	32.82	89.15	25.33
Number additional Rs	1.73	2.30	3.93	4.13	1.07	1.73	.20	.54

The results reported above are singularly unspectacular in that they primarily confirm a host of previous studies of intellectual functioning: the performance of normals is superior to that of schizophrenics while the performance of schizophrenics tends to be superior to that of the brain injured. Further, they fail to demonstrate any significant differences between Process and Reactives. The predictions of the present study, however, concerned differences in the types of *errors* or Inadequate responses made by these groups and the similarity between brain injured and Process subjects in terms of these errors.

As expected, Restrictive errors were more frequent in the Process and brain injured subjects than in the Reactives. This was true not only for mean number of total Restrictive errors but for each of the four separate Restrictive categories; in no case was the mean of the Reactives greater than that of the Process or the brain injured.[3] Exactly the opposite pattern was found for the Expansive errors, the Reactive subjects making more errors of this type than the other two groups. Since the groups were not equal in the total number of Inadequate responses, percentages were computed for each subject to reflect the proportion of total errors that fell into the Restrictive category. Inspection of the group means indicates that the majority of errors made by the Process and brain injured subjects were of the Restrictive type while for the Reactives, less than 50% of their errors were of this type. When these percentage values were compared by means of *t* tests, it was found that the Reactives differed significantly from each of the two other groups ($p < .01$). Although the brain injured made proportionately more Restrictive errors than the Process schizophrenics, the difference between groups was not significant ($t = .76$).

No specific predictions had been developed for the normal controls except the general expectation that they might show less imbalance in type of error than any of the other groups. The results indicated that the performance of these subjects was most similar to the Reactives. That is, the normal subjects made proportionately more Expansive than Restrictive errors. Second, even though the normal subjects made significantly *fewer* total errors (Inadequate responses) than the brain injured and Process subjects, they made significantly *more* Expansive errors, in absolute numbers, than these groups ($p \leq .05$). There were differences, however, between normals and Reactives. The normals, it will be recalled, also gave significantly fewer Inadequate responses than the Reactives. This difference can be accounted for by the greater frequency of Expansive

[3] Although all four categories differentiated the Process and brain injured groups from the Reactives, responses in the Chaining-Object Naming and Concrete Fabulization categories occurred infrequently even in these groups. The most common Restrictive errors were of the Split-narrow and No Response types. Within the two types of Expansive errors, all groups made more Syncretistic and Symbolic responses than Abstract Fabulizations. The Reactives, however, made a fair number of the latter kind of error ($M = 1.53$) although they occurred rarely in the other groups.

errors in the Reactive group. That is, the number of Restrictive errors made by Reactives and normals was approximately equal (*M*s of 1.67 and 1.60, respectively, $t = .10$); the difference between the groups in Expansive errors was, in contrast, significant ($t = 2.55$, $p < .05$). Qualitative factors also distinguished between groups, responses that could be considered as bizarre occurring only in the Reactive schizophrenics.

To emphasize the differences among groups in type of error, a second kind of analysis of the data was undertaken. For each subject the number of Expansive errors was subtracted from the number of Restrictive errors and the sign of the resulting difference score noted. By this simple, if somewhat arbitrary procedure, 13 of the 15 Process subjects were correctly identified (i.e., had the predicted sign), 14 of the brain injured, and 12 of the Reactives. In the normal group, 9 of the 15 subjects gave more Expansive errors and 2 subjects an equal number of each type. A chi square for the first three groups was highly significant ($\chi^2 = 22.20$, with $df = 2$, $p < .01$).

All of the analyses reported above were based on a scoring system in which each of the subject's responses to the 12 groups of objects received only a single score, i.e., was assigned to only one category. Such a system does not take into account the fact that subjects may offer a number of different reasons as to why a particular group of objects belong together. An attempt was therefore made to compare the groups with respect to the tendency to give multiple responses. In order to have a measure uncontaminated by the differential frequency of response failures among the groups, each subject was first given a base score of one point for each item, i.e., no response at all and any single basis of commonality were each given one point. Then the number of additional bases specified in each answer were counted. The means of the total number of additional responses for the 12 items were then determined for each group. As reported in Table 1, it was the Reactive group that was most prone to give additional responses, followed by the normals, the Process, and as a poor fourth, the brain injured. Except for the comparison of the normal and Process groups, the results of t tests indicated that the differences among the groups were significant or approached significance ($p \leqq .10$). Examination of the content of the additional responses indicated that they tended to be of the type predicted for the subject's group. Since the absolute numbers were small, no statistical tests were undertaken.

DISCUSSION

The results of the present study indicate similarities and differences among the groups in terms of both overall "goodness" of performance and types of errors. These results can perhaps best be summarized by attempting to characterize the performance of each group.

In most of their Inadequate responses, the brain injured subjects erred in the direction of being unable to specify a common property that characterized all of the objects. That is, they failed to suggest any basis of commonality, specified a property of only some of the objects in a group, or merely described them in some fashion. However inadequate their responses often were, their approach to the task might at least be characterized as appropriate in that they tended to stay within the limits of the task as specified by the experimenter's instructions and confined themselves in their responses to the fairly immediate properties of the objects.

The performance of the Process schizophrenics was similar to the brain injured in that the predominant types of error made by these subjects were of the same kind. However, their inability to specify common properties of the objects was less marked than the brain injured. That is, they gave significantly more Adequate responses and more often suggested multiple bases for the objects belonging together. Further, for those responses that were considered Adequate, the Process schizophrenics gave more responses that were categorized as Abstract.

Like the Process subjects, the Reactive schizophrenics gave significantly more Adequate responses than the brain injured, and within this category, proportionately more Abstract responses. In both of these measures, however, the two schizophrenic groups

did not differ significantly and thus may be said to have been equivalent in overall goodness of performance. There were marked differences between the groups, however, in the kind of Inadequate responses given. The Reactives might be described as erring in the opposite direction from the Process and brain injured. The Reactives were typically able to suggest some basis for the objects belonging together (response failures were rare). Their responses, however, were frequently overinclusive in the properties they specified, loosely associated with the specific objects, or so bizarre and idiosyncratic as to have no discernible connection with the objects. The willingness of the Reactives to offer a response, however inappropriate, is further demonstrated by the finding that they gave significantly more multiple responses than any other group, including the normals. This frequently had the effect of "saying too much," since within a single answer, it was not unusual to find a high level, abstract response among responses of poorer quality.

The overall performance of the normal control group was, of course, superior to all of the other groups, i.e., they gave a significantly greater number of Adequate responses and a higher percentage of Abstract responses. Examination of the Inadequate responses of this group revealed less of an imbalance between Restrictive and Expansive errors than in any of the other groups. The majority of their errors, however, were of the Expansive type and in this sense, they were more similar to the Reactive schizophrenics than to the other groups.

The results of the present study thus confirmed the prediction that on a conceptual task such as the Goldstein-Scheerer, Process schizophrenics would be more similar in their errors to the brain injured than to Reactive schizophrenics. These findings, in conjunction with the results of Brackbill and Fine (1956) with the Rorschach, add support to the hypothesis that some type of brain pathology may be responsible for the symptoms exhibited by Process schizophrenics, as opposed to etiological factors of a more functional nature in Reactive schizophrenics. However, it hardly seems necessary to state that these results are merely suggestive. Similarities between Process schizophrenics and the brain injured, particularly in behavioral studies such as the present one, may be coincidental in the sense that it is not unusual for disorders of very different etiologies to exhibit overlapping symptom pictures. The results of an investigation by Farina (1960), for example, might be cited as apparently contradictory to the organicity hypothesis. Farina, after observing the interactions of the mother and father of a number of male schizophrenics in his experimental situation, reported that the mothers of Process schizophrenics tended to dominate over their husbands, and that the opposite pattern of interaction occurred in the parents of Reactives. Such findings would seem to be more favorable to functional explanations of both Reactive and Process schizophrenia.

Whether or not the results of the present study are regarded as suggestive of etiological differences between Reactive and Process schizophrenics, they are interesting in other respects. Systems of classifying various "types" of schizophrenia, such as the Kraepelinian one, typically are based on qualitative distinctions between individuals (e.g., different constellations of symptoms). In contrast, the Process-Reactive classification (in practice, often the extremes of a scoring continuum) approaches a quantitative distinction, one group exhibiting deviate behavior earlier, more chronically—and recovering more slowly—than the other. Studies that merely indicate that normals perform more effectively than Reactives and Reactives more effectively than Process do little to remove the suspicion that one may be dealing with one group that is sick and another group that may simply be sicker. Demonstrations of qualitative differences in behavior between Process and Reactives, or of deviations from normal subjects in different directions, as in the present study, give added confidence to the view that there may be more essential differences between the groups than purely quantitative ones.

SUMMARY

Groups of normal, brain injured, and Process and Reactive schizophrenic subjects were given the Goldstein-Scheerer Sorting

test in which subjects were required to specify why groups of common objects "belong together." These groups, each of 15 subjects, were matched on age and education. Responses that were considered inadequate were assigned to one of two error categories, one reflecting the inability of the subject to specify a common property of all of the objects and the second the specification of properties that were too broad, vague, or idiosyncratic. It was predicted that the errors of Process schizophrenics and brain injured subjects would predominantly be of the first type, the errors of Reactive schizophrenics predominantly of the second type, and that normals would exhibit less of an imbalance in type of error than the other groups. The results confirmed these predictions. In addition, it was found that normals were significantly superior to the other groups in the number and the quality of their adequate responses. The two schizophrenic groups were superior to the brain injured on these latter measures, but did not differ significantly from each other.

REFERENCES

BELLAK, L. The schizophrenic syndrome: A further elaboration of the unified theory of schizophrenia. In L. Bellak (Ed.), *Schizophrenia: A review of the syndrome*. New York: Logos, 1958. Pp. 216–278.

BOLLES, M., & GOLDSTEIN, K. Impairment of "abstract behavior" in schizophrenic patients. *Psychiat. Quart.*, 1938, **12**, 42–65.

BRACKBILL, G. A., & FINE, H. J. Schizophrenia and central nervous system pathology. *J. abnorm. soc. Psychol.*, 1956, **52**, 310–313.

CAMERON, N. Reasoning, regression, and communication in schizophrenics. *Psychol. Monogr.*, 1938, **50**(1, Whole No. 221).

FARINA, A. Patterns of role dominance and conflict in parents of schizophrenic patients. *J. abnorm. soc. Psychol.*, 1960, **61**, 31–38.

GOLDSTEIN, K., & SCHEERER, M. Abstract and concrete behavior: An experimental study with special tests. *Psychol. Monogr.*, 1941, **53**(2, Whole No. 239).

HARRIS, J. G., JR. Size estimation of pictures as a function of thematic content for schizophrenic and normal subjects. *J. Pers.*, 1957, **25**, 651–671.

MCDONOUGH, J. M. Critical flicker frequency and the spiral aftereffect with process and reactive schizophrenics. *J. consult. Psychol.*, 1960, **24**, 150–155.

MCGAUGHRAN, L. S., & MORAN, L. J. "Conceptual level" vs. "conceptual area" analysis of object sorting behavior of schizophrenic and nonpsychiatric groups. *J. abnorm. soc. Psychol.*, 1956, **52**, 43–50.

PHILLIPS, L. Case history data and prognosis in schizophrenia. *J. nerv. ment. Dis.*, 1953, **117**, 515–525.

RAPAPORT, D. *Diagnostic psychological testing* Vol. 1. Chicago: Yearbook Publishers, 1945.

RODNICK, E. H., & GARMEZY, N. An experimental approach to the study of motivation in schizophrenia. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1957*. Lincoln: Univer. Nebraska Press, 1957. Pp. 109–184.

ZAHN, T. P. Acquired and symbolic affective value as determinants of size estimation in schizophrenic and normal subjects. *J. abnorm. soc. Psychol.*, 1959, **58**, 39–47.

(Received September 5, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 395–402

THE EFFECTIVENESS OF "OVERHEARD" PERSUASIVE COMMUNICATIONS

ELAINE WALSTER AND LEON FESTINGER [1]

Stanford University

It is widely believed that a communication, if inadvertently overheard, is more likely to be effective in changing the opinion of the listener than if it had been deliberately addressed to him. Intuitively, the advantages of an overhearing situation for effective influence seem so obvious that it is accepted as virtually proven. Berelson (1950), for example, gives two reasons for the presumed extraordinary effectiveness of overheard communications. He states that,

> in such exposure defenses against new ideas are presumably weaker because preconceptions are not so pervasively present. Finally, there may be other psychological advantages centering around the gratification of "overhearing" something "not meant for you," a consideration that also weakens the resistance to "propaganda" (since "it would not be propaganda if it wasn't intended for you") (p. 458).

This last parenthetical remark is amplified by Hovland, Janis, and Kelley (1953) as follows:

> Remarks such as those overheard in subways and other crowded public places would be especially effective in this respect because under such circumstances it is quite apparent that the speaker has no intention to persuade the bystanders (p. 23).

In short, on an intuitive level, there seem to be three factors which increase the effectiveness of overheard persuasive communications.

1. Because of the accidental nature of the communication the listener is caught, so to speak, with his defenses down.

2. It is more powerful because the listener is not supposed to hear it.

3. The speaker cannot possibly be seen as intending to influence the listener.

Let us see to what extent there is experimental evidence to support these intuitively plausible assertions. On the first factor, namely, lack of defenses when the communication is unanticipated, there are some supporting experimental data. It has been shown (for example, Allyn & Festinger, 1961; Ewing, 1942) that if a person expects a communication to disagree with the opinion he holds, he does erect defenses, tends to reject the communicator, and is less influenced than if he hears the same communication without having been led to expect disagreement. It should be pointed out, however, that these studies did not involve accidental communications. There is no evidence in the literature that there would be fewer defenses to an accidental communication than to a purposeful one, provided the listener did not expect disagreement.

On the second factor, the gratification of hearing something one is not supposed to hear, there is no evidence at all. To our knowledge no studies which bear on this have been done.

On the third factor, however, there is some evidence. There are studies in the literature which have investigated the effectiveness of communications from biased and unbiased sources. A biased source may be seen as having some ulterior or personal motive for wanting to influence the listener. Hence, the difference in effectiveness of communications from biased and unbiased sources may be seen as relevant to the factor of whether or not the speaker is seen as "intending" to influence the listener. The evidence on this point, however, is ambiguous. Kelman and Hovland (1953), for example, find that when a communication is ascribed to a selfishly motivated communicator there is less influence. Hovland and Mandell (1952), however, find no difference in the amount of opinion change when the communication is ascribed to a partial or impartial person. This latter finding is of particular interest since the lack of difference in effectiveness existed in spite of the fact that the listeners rated

[1] The authors would like to express their appreciation to the Ford Foundation for its support of this research.

the impartial source as having done a better job and having been more fair and honest.

Clearly, the evidence for the assertion that overheard communications are especially effective in changing opinions is less than adequate. The first of the two studies to be reported in this paper was designed to provide some better evidence on this point. In order to narrow the field of possible interpretations, the study was designed to avoid entirely any differences which might be due to direct versus indirect persuasive communications. Thus, two conditions were set up in the experiment. In both conditions a subject listened to two persons carrying on a conversation. The only difference between the two conditions was that in one the subject thought that the speakers knew that some people were listening, while in the other the subject thought that the speakers were unaware that anyone at all was listening. The exact details of the experiment follow.

EXPERIMENT I

Procedure

During a lecture in the introductory psychology course at Stanford University, students were informed that, as part of the course, they were required to sign up for a 15-minute tour of the social psychology laboratory's observation room. The next day, in one-half of the recitation sections, arrangements were made with students to tour in groups of 2–4. Students in the other sections were told arrangements would be made with them later in the term—these students were then utilized as a control group. One hundred subjects actually took the tour; 84 subjects comprised the control group.

When subjects reported for the tour, the mechanics and common uses of the observation room (a small booth from which one could look into a larger room through a one-way mirror and could listen through earphones) were explained. Subjects were told that a common technique for analyzing group behavior was "blind listening." The experimenter explained that they would be told more about, and given a chance to try, this technique later. After these instructions the experimenter took the subjects to the observation room.

As the groups began walking down the corridor toward the observation room, the experimenter remarked casually,

> Since the large observation room is rarely in use for experiments, graduate students usually use it as a lounge. There's almost always someone in there.

One-half of the groups were given instructions to convince them the graduate students would know they were listening; the other subjects were led to believe the graduate students would be unaware of their presence.

For the Regular condition, subjects were told that,

> A couple of graduate students are in there now. I'll tell them we're coming through.

Upon reaching the end of the corridor, the experimenter stuck her head into the large room to the right of the small observation booth, and said to the two graduate students within,

> Hi. Some Psych I people are coming through. Would you mind just continuing your conversation so we can try a little blind listening?

For the Overheard condition, the experimenter said instead,

> I think a couple of graduate students are in there now. If we're quiet, we can probably get into the observation booth without their hearing us.

The experimenter then led the subjects quietly past the large room into the booth, shut the door of the booth, and looking through the one-way mirror at the graduate students, who had not looked up, said,

> Well, I guess we made it.

In both conditions, as subjects watched the graduate students talking, the experimenter described some of the techniques of blind listening. (Noticing incomplete sentences, pauses, etc.) Subjects were then told to try a little blind listening. Curtains were drawn over the one-way mirror, and earphones passed out. (Instruction in this "technique" had been so sketchy and uninformative that subjects really had nothing to do but listen casually.)

The "conversation" all subjects heard through the earphones was actually a taped 6-minute persuasive communication. On this tape, two speakers discused the common "misconception" that smoking causes lung cancer. A series of (nonexistent) studies were described which showed there was no relationship between smoking and cancer when all confounding factors were eliminated. In fact, the data were said to suggest that smoking might even be beneficial since it released tension.

The degree to which the subjects agreed with the communication was assessed a week later. A mimeographed questionnaire purporting to be from the National Institutes of Health was given to the 100 experimental and the 84 control subjects during the regular class recitation sections by a second experimenter. The "medical survey" included four questions concerning students' opinions about the causal relationship between smoking and cancer.

About 10 minutes after this health questionnaire was completed, the first experimenter entered the classroom and passed out a second questionnaire asking students to evaluate the tour she had con-

ducted. These data were presumably "to be used in planning next quarter's demonstrations."

Fourteen of the 100 subjects in the experimental conditions indicated on this second questionnaire that they thought the conversation they heard from the observation booth was a prepared recorded conversation. The data from these 14 subjects were discarded.

Following the collection of data, the true purposes of the experiment were explained to all subjects.

Results

The answers to the four questions on the Health Questionnaire which concerned their opinions about the link between smoking and lung cancer provided a measure of the relative extent to which the persons in our various conditions were influenced by the persuasive communication they listened to in the guise of a conversation. Since all the questions yielded comparable results, we summed the responses to all four questions to provide one measure of agreement with the communication. On each question an answer was given a higher score the more it agreed with the position that the link between smoking and lung cancer was *not* proven—the position advocated by the communication. Table 1 presents these data for the three conditions.

It is clear from an examination of Table 1 that the conversation between the two graduate students was an effective persuasive communication. While the control group (who did not hear the conversation) shows an opinion score of 11.4, those in the Regular condition have a score of 13.6, and those in the Overheard condition have a score of 14.5. The differences between the Control condition and each of the experimental groups are both highly significant. We are, however, more concerned with the difference between the two experimental groups. As seemed plausible and as most persons would have expected, the Overheard condition produces more opinion change than the Regular condition. The difference is significant at the .08 level of significance ($t = 1.76$). In other words, if a person feels that the speaker does not know that anyone is listening he is influenced more than if he thinks the speaker knows someone is listening.

TABLE 1

OPINIONS CONCERNING THE LINK BETWEEN SMOKING AND LUNG CANCER

Condition	Smokers	Nonsmokers	Total
Students not participating in the demonstration (Controls)	11.8 (24)[a]	11.2 (60)	11.4 (84)
Regular	13.6 (12)	13.6 (36)	13.6 (48)
Overheard	15.3 (9)	14.2 (29)	14.5 (38)

[a] Number in parentheses is the number of cases on which the cell mean is based.

A difference which is of such borderline significance statistically, however, is not entirely satisfactory. We, hence, attempted to gain further insight into the data by doing some further analyses. Because of the content of the persuasive communication, the first variable which suggested itself for internal analysis was, of course, the smoking behavior of the subjects in the study. It seemed possible to us that, since smokers might be very desirous of believing the content of the persuasive communication, they might have changed a great deal, and equally, in both experimental conditions thus, perhaps, attenuating the difference. The analysis of the data, however, revealed a very perplexing finding. Not only was this suggestion not true, but the opposite was true. Table 1 presents the data separately for smokers and nonsmokers for the three conditions.

It is clear in this table that the differential effectiveness of the Regular and Overheard conditions is due almost entirely to the small group of smokers. For the nonsmokers there is a very small difference between these two conditions which is not significant. For the smokers, however, the difference is significant at better than the .10 level ($t = 1.88$) even though the number of cases is very small. It is also clear that the smokers are not more influenced than are the nonsmokers in the Regular condition. These means are almost identical. It is only in the Overheard condition that the smokers, who presumably would like to believe the content of the communication, are actually more influenced than the nonsmokers.

In an attempt to understand the reasons for this surprising finding we looked at the

data from the second questionnaire where the subjects were asked how sincere and honest they thought the speakers had been. Perhaps the smokers, for some reason or other, attributed greater sincerity to the speakers in the Overheard condition than did the nonsmokers. The data show, however, that this cannot account for the obtained result. Everyone, it turns out, rates the speakers as more honest and sincere in the Overheard than in the Regular condition. In the Overheard condition the mean ratings are 6.4 and 6.1, for the smokers and nonsmokers, respectively. The comparable figures for the Regular condition are 5.6 and 4.8. Thus, actually, the nonsmokers are the ones with the bigger differential between the two conditions in the rating of the speakers' sincerity but it is the smokers who show a differential in how influenced they were. None of the other questions that were asked on this second questionnaire revealed any differences at all in reactions to the two conditions.

The remaining alternative is that the obtained result may be a chance difference. After all, it seems somewhat implausible that the Overheard condition should only be more effective for persons who are behaviorally involved in the issue. We, consequently, carried out a second experiment which was more adequately controlled and which was specifically designed to see if this implausible finding was or was not correct.

EXPERIMENT II

Procedure

Two groups of persons, wives of college students living in an on-campus housing development, and Junior and Senior women students living in dormitories, were used as subjects in the experiment. Two persuasive communications were employed. One of them, "Student husbands should spend a great deal more time at home," was expected to be an involving issue primarily for the married women. The other, "Junior and Senior women should be allowed to live off campus if they desire," was relevant only for the unmarried students living in dormitories.

Two weeks before the experimental sessions, prospective subjects were sent letters from the National Institutes of Mental Research, informing them that a nationwide study of the therapy process was being conducted, and inviting them to observe and evaluate a regular therapy session. Final appointments were made with 41 married women and 40 single students in follow-up telephone calls. Groups of one to four subjects were scheduled at a time. One-half of these groups were randomly assigned to the Regular condition; the remaining groups were placed in the Overheard condition.

After all the women in any group had arrived, the supposed "purposes" of our study were discussed at greater length. Subjects were told they were to evaluate two therapists, the patient's regular therapist and a visiting psychiatrist who was consulting with the patient's regular therapist. Subjects were instructed just to listen casually to the therapy session in order to give us some very general impressions at the end. It was explained to the subjects that they were to listen to the sessions from individual listening rooms connected to the therapy room by a one-way speaker. Separate listening rooms were used "since we wanted independent judgments" and since they "obviously could not sit *in* the therapy room."

At this point Regular condition subjects were told,

> Of course both the therapists and the patient know you'll be listening in today.

Those in the Overheard condition were told,

> Of course, the therapists don't know you're listening in today.

As they were led to their individual rooms, the experimenter commented to Regular condition subjects,

> Now I'll run downstairs and tell the therapists you're here and turn on the main speaker.

Overheard condition subjects were merely told,

> You'll be able to hear as soon as I turn on the main speaker.

Instead of "turning on the main speaker," the experimenter actually turned on a tape recorder which began playing a casual conversation between the two "therapists," through the subjects' speakers. The experimenter then returned to the subjects' rooms, identified the therapists' voices, and remarked to the subjects in the Regular condition,

> When I looked in downstairs, Bill [the patient] hadn't arrived yet.

Subjects in the Overheard condition were told,

> Bill apparently isn't there yet.

The experimenter then left the subjects alone to hear the remainder of the tape. The "conversation" between the two waiting therapists continued for 14 minutes. During his discussion of student problems, the visiting therapist spent 6 minutes delivering the first persuasive communication, "Women should be allowed to live off campus." The regular therapist only commented in order to clarify and agree with the arguments of the visitor. The regular

TABLE 2

INDEX OF AGREEMENT ON RELEVANT AND IRRELEVANT ISSUES FOR MARRIED WOMEN AND SINGLE STUDENTS

Issue	Married women		Single students	
	N	Mean index score	*N*	Mean index score
Involving				
Control subjects	18	7.11[a]	18	7.57[b]
Regular condition	20	6.71 $p = .01, t = 5.18$	20	8.85 $p = .02, t = 2.45$
Overheard condition	21	8.48	20	9.69
Noninvolving				
Control subjects	18	7.20[b]	18	8.58[a]
Regular condition	20	8.47 *ns*	20	8.17 *ns*
Overheard condition	21	8.75	20	8.49

[a] Husband issue.
[b] Dormitory issue.

therapist then began discussing a problem "common to his patients." During the next 6 minutes, he presented the second persuasive communication, "Husbands should spend more time at home."

At the end of this discussion, the patient arrived, his therapy session began, and for the next 35 minutes both therapists limited their remarks to "Yes," "No," and extremely obvious interpretations. At the end of the 50-minute taped "session," subjects were asked for their evaluations of the therapists' competence—supposedly the focus of our study. The experimenter then gave them a questionnaire asking for their opinions on several problems of concern to Stanford students and student-wives. From a few of the questions in this attitude survey, an Index of Agreement with each of the communications presented by the therapists was computed.

Control Group. In order to get some indication of the attitudes of subjects who did not hear the tape, the questionnaire was also mailed to 21 married women and 20 single students selected in the same way as our experimental subjects. Seventeen married women and 17 students returned the completed forms. Since Control subjects answered the questionnaire at home with unlimited time to complete it, their responses may not be completely comparable to those of our experimental groups.

Results

Each subject in the experiment heard two persuasive communications, one of which was supposed to be personally involving and the other not. It is important to assess the extent to which we were successful in accomplishing our purpose in the design. On the questionnaire, subjects were asked how intensely they felt about each of the issues and also how often they discussed each of them. The results show that the appropriate conditions were, indeed, created. The dormitory women report feeling more intensely about the "off-campus living" issue than do married women ($t = 3.98$) and report having discussed it much more often ($t = 8.32$). On the "husbands at home" issue, on the other hand, the married women feel more intensely than do the dormitory women (although not significantly so) and report discussing the issue much more often ($t = 8.05$).

We may then look at the results to see if the findings from the first experiment do or do not hold up. In order to present one measure, for each issue, of extent of agreement with the persuasive communication, the answers to the relevant questions on the questionnaire were summed. In all cases higher scores were assigned to those responses which indicated more agreement with the position advocated in the communication. Table 2 presents these data.

An examination of the data makes it clear that we obtained precisely the same kinds of results in this experiment as were obtained in the first experiment. When the issue is one which is immediately involving for the subject ("husbands at home" for married women and "living off campus" for dormitory students) there are large and significant differences in the persuasive effectiveness of the communication between the Regular and Overheard conditions. When the issue is not immediately involving, the differences between the Regular and Overheard conditions are negligible and not significant.

For the dormitory issue, the interaction between listening condition (Overheard or

Regular) and degree of subjects' involvement, was significant at $p < .05$ ($F = 5.44$, $df = 1/77$). The interaction was not significant for the "Husbands at home issue."

It is worth emphasizing at this point that the design used in this experiment is rather well controlled. The types of subjects and content of issues were chosen for the specific purpose of testing the suggestion from the first experiment and each subject is, in a sense, her own control. The same subjects who are more affected by the "overheard" conversation when the issue involves them actively are *not* more affected by the "overheard" communication when the issue does not involve them. The consistency of the results we have obtained warrants rather high confidence in the effect. The degree to which the results are consistent may be seen by looking simultaneously at the results of both experiments. We have presented these results graphically in Figure 1.

The data in Figure 1 are plotted, for all three issues used in the two experiments, in terms of presumed change of opinion as a result of hearing the persuasive communication. In other words, the data are plotted in terms of differences from the control groups who never heard the communication. The solid lines indicate data from subjects for whom the particular issue was an actively involving one. The dashed lines represent the data when the issue was not involving for the subjects. The consistency of the data is so immediately apparent in the figure that it hardly needs comment. We may certainly conclude that, at least when the position advocated by the persuasive communication is one which the involved subjects would like to accept, overheard communications are more effective for involved subjects and not particularly more effective for subjects not directly involved in the issue.

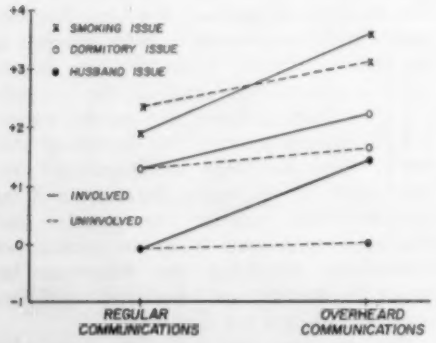

FIG. 1. Opinion change with "regular" and "overheard" communications.

DISCUSSION

There is clearly a problem involved in explaining these results. Let us say right now that we have no explanation to propose in any true sense of a theoretical explanation. There are, however, some specific directions into which one seems to be forced by the data, and these are worth discussing.

We will begin this discussion by examining what our results have to say about the three "plausible" factors which have been generally presumed to make overheard communications more effective. It will be recalled that we listed these and discussed them somewhat in the introduction to this article. We will examine them in relation to our data in turn.

1. Because of the accidental nature of the communication the listeners' defenses are not prepared. This factor, whether or not it is relevant to other situations, cannot conceivably account for our data. After all, we were not comparing a "direct attempt to influence" with an "overheard communication." Neither of our experimental conditions represented direct attempts to influence the subject. In neither case were they prepared for the communication. There is also another interesting point that should be made about the meaning of the term "defenses," a term widely used in discussing resistance to persuasive communications. If one does a study in which one attempts to persuade persons of something which they do not want to believe, it is intuitively plausible to think of the person defending himself. But in our study we were not doing this. Why would a smoker defend himself against the opinion that smoking is not harmful; why would a student living in a dormitory defend herself against the opinion that, if she wanted to, she should be allowed to live off campus?

2. Because the listener is not supposed to hear it, an overheard communication is

more powerful. This factor, at least as specified, also cannot account for our results. In the overheard conditions neither the subjects who were or who were not involved in the issue were "supposed to hear the communication." Nevertheless, we obtain a difference favoring the Overheard condition only when the issue is one which is involving for the subjects. One might argue that perhaps the psychological interest aroused by hearing something one is not intended to hear is greater if the content is immediately relevant. This, of course, is possible. But, if this is true, it would seem equally plausible that psychological interest would be heightened equally in the Regular condition if the content is immediately relevant. Consequently, it is difficult to see why this factor alone would make for a differential between our two conditions.

3. Because the speaker does not know the listener is there, the speaker cannot be seen as intending to influence the listener. This factor seems to be the one to which we are driven in our search for an explanation of our findings. It seems reasonable that if a speaker is seen as intending to influence one, then suspicions concerning possible ulterior motives may serve to nullify the possible effectiveness of the communication. But how does this apply to our data? Let us look at our experimental conditions and examine to what extent it would be possible, in any of them, to impute "intention to influence" to the speaker. First of all, let us dispose of the overheard conditions. Clearly, since the speaker does not know anyone is listening, it is not conceivable that the listener, knowing this, could imagine that the speaker intends to influence him. In the Regular condition, however, the situation is a little different. Here the speaker knows that someone is listening but does not know who. Since the listener in the Regular condition knows he has not been personally identified by the speaker, if the content of the communication is not personally relevant, it would also seem quite difficult to imagine that the speaker was intending to influence him. If, however, the content of the communication in the Regular condition is of direct personal relevance, then it is possible for the listener to feel that the speaker is, indeed, attempting to influence him and has ulterior motives.

If this is correct, then one comes to the conclusion that the difference between the two conditions for those who were involved in the issue exists because the effectiveness of the communication in the Regular condition was partially nullified by the possibility of imputing ulterior motives to the speaker. When this possibility is eliminated, as in the Overheard condition, those involved in the issue change more in the direction of the communication. This is, of course, to be expected since the communication urges them in a direction they would like to move.

There is, of course, still the interesting question, to which we have no answer, as to why, when the involved subjects hear a communication *which advocates a position they would like to accept*, they still tend to impute ulterior motives to the speaker when this is at all possible.

Clearly, these are suggestions for explanation and for futher exploration rather than actual theoretical explanation of the findings. The results certainly indicate that the effectiveness of overheard communications is a more complicated matter than has been generally assumed. If imputation of ulterior motives to the speaker only comes into play, or comes primarily into play, when the content of the communication is personally involving, this may indeed account for the inconsistency of previous experimental results on this topic. And it may also be, of course, that if the communication urges the listeners in a direction in which they do not want to move, quite different results might be obtained.

SUMMARY

Two experiments were conducted to assess the effectiveness of overheard communications:

1. When subjects touring an observation room "overheard" a persuasive communication, they changed their attitudes more than when they were told speakers knew they were listening. Quite unexpectedly, those subjects for whom the communication (smoking is not harmful) was especially relevant and involving (smokers) accounted almost entirely for the difference between the two conditions.

2. A second study was done to verify this finding. Two groups, married women and women in dormitories, were each given two communications, one concerning husbands and one dormitory living. It was expected that when the issue was an involving one for the group, the overheard communication would be especially effective. When the issue was not involving, the difference between the overheard and regular conditions was expected to be small. The results completely confirmed those expectations.

REFERENCES

ALLYN, JANE, & FESTINGER, L. The effectiveness of unanticipated persuasive communications. *J. abnorm. soc. Psychol.*, 1961, **62**, 35–40.

BERELSON, B. Communication and public opinion. In B. Berelson & M. Janowitz (Eds.), *Reader in public opinion and communication*. Glencoe, Ill.: Free Press, 1950. P. 458.

EWING, T. A study of certain factors involved in changes of opinion. *J. soc. Psychol.*, 1942, **16**, 63–88.

HOVLAND, C. I., JANIS, I. L., & KELLEY, H. H. *Communication and persuasion*. New Haven: Yale Univer. Press, 1953.

HOVLAND, C. I., & MANDELL, W. An experimental comparison of conclusion drawing by the communicator and by the audience. *J. abnorm. soc. Psychol.*, 1952, **47**, 581–588.

KELMAN, H. C., & HOVLAND, C. I. "Reinstatement" of the communicator in delayed measurement of opinion change. *J. abnorm. soc. Psychol.*, 1953, **48**, 327–335.

(Received September 18, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 403–410

AN EXPERIMENTAL STUDY OF PSYCHOLOGICAL RESISTANCES TO FEAR AROUSING COMMUNICATIONS [1]

IRVING L. JANIS
Yale University

AND

ROBERT F. TERWILLIGER
Douglass College

As part of the Yale program of research on attitude change, a series of experimental studies (Janis & Feshbach, 1953, 1954; Janis & Milholland, 1954) have been carried out for the purpose of determining the conditions under which fear arousing communications are effective or ineffective. The present study was designed to follow up on the implications of the earlier findings concerning defensive reactions which interfere with acceptance of the communicator's message.

The initial experiment in the series (Janis & Feshbach, 1953) consisted of exposing equivalent groups of high school students to three forms of an illustrated lecture on dental hygiene, which varied in the amount of fear arousing material presented. The "strong threat" version, which emphasized the threat of pain, disease, and body damage, was found to be less successful than the "minimal threat" version in producing acceptance of the communicator's recommendations. These results suggested that a communication will produce more resistance to attitude change if it arouses a high level of fear than if it arouses a lower level of fear. A subsequent experiment by Haeffner (1956), comparing the effectiveness of high and low threat versions of a persuasive communication concerning the desirability of international control of nuclear weapons in order to prevent a devastating atomic war, has provided some confirmatory evidence in support of this hypothesis.[2]

Taking account of additional findings on predispositional differences, as well as the experimental results based on the comparisons of groups exposed to mild versus strong threat appeals, Janis and Feshbach (1954) have formulated the following theoretical analysis:

> We can view fear appeals as stimuli which are capable of arousing an anxiety drive in the recipients. This drive state motivates the individual to engage in many different types of internal responses that are oriented toward the goal of escaping from anxiety-arousing cues. A major type of response that would tend to facilitate acceptance of the communicator's message consists of anticipations of averting the threat depicted in the communication through con-

[1] This experiment was conducted in connection with an undergraduate research stipend awarded by the Social Science Research Council to the junior author in 1956, and was carried out under the auspices of the Yale Communication and Attitude Research Program, which is supported by a grant from the Rockefeller Foundation. The authors wish to thank the late Carl I. Hovland for his valuable suggestions and criticisms.

[2] Two attempted replications of the Janis and Feshbach experiment have also been reported, one with newly inducted recruits at a United States Air Force Base (Moltz & Thistlethwaite, 1955) and the other with students at a high school in the State of Washington (Goldstein, 1959). Both studies used a dental hygiene communication similar to the one used by Janis and Feshbach, but modified it by presenting all the threatening material first and then presenting all the reassurances and recommendations at the end. Unfortunately, the evidence from both experiments indicates that with their modified versions of the communication, the experimenters did not succeed in their attempts to manipulate the degree of fear arousal. Moltz and Thistlethwaite report data on postcommunication dental "anxiety" scores which show that "the communication materials intended to be differentially fear arousing were not effective in producing differences in anxiety arousal" (p. 236). Goldstein presents no relevant data on this question in his published report (1959); but in a more complete manuscript (1957) one finds that he also obtained no significant differences on immediate postcommunication measures of dental "anxiety." Consequently, neither of the attempted replications can be regarded as providing an adequate test of any hypothesis concerning differences in amount of attitude change induced by different degrees of fear or anxiety arousal. The experimenters, nevertheless, have published their attitude change findings concerning conformity to the dental hygiene recommendations: One study (Moltz & Thistlethwaite) reports no significant differences; the other (Goldstein) reports a difference between the strong and mild fear appeal groups (without presenting the significance test data) and claims that it supports Janis and Feshbach's conclusion concerning the greater effectiveness of a minimal fear appeal (p. 251).

formity with the suggested recommendations. These facilitating responses probably are most likely to occur under the following conditions: (a) When a relatively low strength of anxiety drive is aroused; (b) when a specific external threat is depicted in the communication and is perceived to be one that could, somehow or other, be averted; and (c) when the communication contains reassuring recommendations which are perceived to be a potentially successful means of averting the threat. When these conditions are not met, however, we assume that the predominant responses will tend to be of an interfering nature—inattentiveness, perceptual distortions, defensive efforts to deny or minimize the threat, etc. (p. 162).

In the previously mentioned experiments, all the findings bearing on defensive reactions were based entirely on changes in verbal opinions expressed on questionnaires given *before* and *after* exposure to the entire communication. No data were obtained on the subject's reactions during *actual exposure* to the communication. It therefore remains an open question whether the subjects' resistances are mobilized at the time when their fear is aroused, as the foregoing theoretical analysis assumes. The present experiment was designed primarily to fill this gap by investigating manifestations of resistance in two groups of subjects during the period when they are reading a Low or High Threat version of a public health communication. The central hypothesis is that during exposure to a persuasive communication, there will be more manifestations of resistance if it includes strong threat statements that arouse a high degree of fear than if it includes milder threat statements.

METHOD

Each subject was interviewed in a private session, during which he was exposed to a communication which fostered the conclusion that heavy smoking causes cancer and which recommended that everyone should avoid or cut down on cigarette smoking. The experiment was designed to assess the immediate reactions evoked by two versions of the communication, one containing a Low Threat appeal and the other, a High Threat appeal. The influence of the High versus Low Threat appeals was investigated by comparing the subjects' spontaneous associations elicited during exposure to the communication and by observing changes in the subjects' responses to a standard interview administered before and after exposure to the communication.

Two Versions of the Communication

The statements presented in both versions of the communication consisted mainly of authentic quotations from prominent medical authorities, supplemented by a few transitional statements written by the experimenters. The subjects were told that this was a scientific digest prepared by members of the faculty of Yale University and that all the quotations were from medical authorities. The two versions of the communication were checked for accuracy by several members of the faculty of the Yale Medical School and by the New Haven chapter of the American Cancer Society.

Both versions included the same set of 15 paragraphs concerning the unfavorable consequences of smoking, which we refer to as the "standard communication." It included three types of material: discussion of the existing evidence on the relationship between smoking and cancer (9 paragraphs); statements about the seriousness of the types of cancer produced by cigarette smoking (3 paragraphs); and reassuring recommendations to the effect that one can avoid the danger by cutting down on smoking (3 paragraphs). The first category included 6 paragraphs devoted to authentic quotations which presented positive evidence indicating that smoking greatly increases one's chances of developing cancer. But, in order to maintain a tone of scientific objectivity, 3 paragraphs were included which mentioned somewhat negative evidence or which criticized some of the positive evidence, conveying the notion that "this outcome might not necessarily be the case" or "cancer is likely to develop only in exceptionally heavy smokers."

The standard communication is regarded as constituting a mild threat appeal because it merely called attention, in highly objective and abstract terms, to the dangerous consequences of cancer. The following is a typical example of the mild way in which the threat statements were worded in the standard communication:

> Since cancer of the lungs is more likely to develop from moderate smoking than is cancer of the lips, some experts have advised all smokers to take account of the symptoms and prognosis of lung cancer. In contrast to cancer of the lips, cancer of the lung is described by medical authorities as a highly malignant disease which is extremely difficult to diagnose before it is too late for effective treatment.

The mild threat version of the communication consisted only of the 15 paragraphs in the standard communication, whereas the strong version included 7 additional paragraphs which were interspersed among the standard ones. These 7 paragraphs said nothing about smoking but elaborated on the statements about the seriousness of lung cancer. Although objective in tone, they emphasized the painful symptoms, the body damage, and the fatalities caused by lung cancer, conveying a much more detailed picture of the dire consequences of the disease. The

following is typical of the strong threat appeal paragraphs:

> American Cancer Society: Anywhere along the respiratory tract, a single cancerous cell can start the growth of a malignant tumor that eventually may kill. Such a tumor can prove fatal either by causing suffocation or by sending deadly cells into other vital parts of the body.
>
> Dr. Alton Ochsner (cancer specialist): Lung cancer necessitates drastic surgery. In only one out of three victims brought to the hospital will the cancer be localized sufficiently so that the cancerous lung can be removed. Of these, 85% will be dead within 5 years, most of them within 2.

Subjects

The experiment was conducted with 31 adult subjects, 18–55 years of age, including 6 females and 25 males. The younger subjects were male college students who were obtained as volunteers from undergraduate classes, while the older subjects were relatively well educated, middle-class men and women who volunteered to participate under the auspices of the local chapter of the American Cancer Society. Subjects were assigned on a stratified random basis to the two experimental groups, so that there were approximately equal numbers at each educational and age level. The High Threat group consisted of 14 subjects, including 5 nonsmokers (36%), 6 light or moderate smokers (43%), and 3 heavy smokers, who consumed a pack or more per day (21%). In the Low Threat group there were 17 subjects, and the distribution of smokers was almost identical: 6 nonsmokers (35%), 7 light or moderate smokers (41%), and 4 heavy smokers (24%).

Experimental Procedures

Each subject was seated in a booth in such a way that both the experimenter and the tape recording equipment were out of sight. At the beginning of the session he was given some preliminary training in verbalizing aloud his spontaneous associations to the sentences in a "practice" communication on the enjoyment of classical music. The instructions requested the subject to read aloud the paragraph on each card and then to give all his thoughts and feelings about the statements he had just read. He was told that he could stop at any time and give his associations but that he should always stop to do so at the end of each paragraph.

The subject was then accustomed to an experimental device, called "the auditory feedback suppressor," which was developed by one of the authors for the purpose of enabling the subject to speak aloud without hearing what he is saying, on the assumption that under these conditions the subject's verbalizations will more closely approximate his own silent thoughts (see Janis, 1959, p. 214). The subject's ears were plugged with cotton and he wore earphones which delivered a constant white noise of sufficient intensity to mask the sound of his own voice. After having become accustomed to the sound itself, the subject was trained to give his associations while the sound was on. For this purpose, the subject was exposed to another short communication dealing with Civil Defense.

Following this training communication, the subject was interviewed concerning his smoking habits. The main attitude question consisted of asking the subject to verbalize his thoughts and feelings in response to the following role playing instructions:

> Now I would like you to concentrate on the idea of yourself smoking and tell me whatever comes to your mind while you are thinking about this. Be sure that you tell me everything you think of when you imagine yourself as taking out a cigarette and smoking it. Give whatever thoughts occur to you, no matter how irrelevant they may seem to you. Please begin when I turn on the noise.
>
> [When the subject pauses after giving his initial associations]: What other thoughts come to mind when you think about smoking? Concentrate again on the idea of yourself smoking and keep telling me whatever comes to your mind.

From the time the interview was begun, the entire session was tape recorded and the auditory feedback suppressor was continuously used so that the subject heard only the constant background noise which prevented him from hearing his own voice.

After the precommunication interview, the subject was told that he would next be given a public health communication that was being tried out before being released as a published pamphlet. He was instructed to follow exactly the same procedure used during the training part of the session with the preliminary communications, the main rule being to give his spontaneous associations at the end of each card. He was then exposed to either the Low Threat or High Threat version of the communication, each paragraph being presented in printed form on a separate card. Following a brief rest period at the end of the communication, the main attitude question was presented for a second time in order to assess changes in the subject's attitude concerning cigarette smoking.

Content Analysis Procedures

The tape recordings were transcribed onto typescript sheets, and the subject's responses were then analyzed by using a systematic content analysis procedure. The spontaneous associations given in response to each paragraph of the communication were scored in terms of the following categories:

A. Affective Reactions

1. Expressions of worry, affective disturbance, or emotional tension evoked by the communication. Example: "That is awful."

2. Reference to the unpleasant aspects of cancer—pain, suffering, body damage. Example: "I never realized what a terrible disease it is." (This category includes paraphrases of those specific

statements in the communication which describes the suffering of cancer victims; however, paraphrases of statements to the effect that cancer is *caused by smoking*, were scored as B_s.)

B. Evaluative Comments

1. Major criticisms, i.e., explicit rejection statements to the effect that a given argument or conclusion concerning the harmful effects of *smoking* is incorrect, unwarranted, or unacceptable. Example: "I don't believe that."

2. Minor criticisms, i.e., unfavorable comments about the style or objectivity of the material being read that imply but do not manifestly assert that the argument is unacceptable. Example: "This argument is not clear."

3. Major favorable comments, i.e., explicit acceptance statements to the effect that a given argument or conclusion concerning the harmful effects of *smoking* is correct, warranted, or in agreement with the subject's own views. Example: "I go along with that."

4. Minor favorable comments, i.e., positive statements about the style or objectivity of the material being read that imply but do not manifestly assert that the argument is acceptable. Example: "That is an impressive way to put it."

5. Paraphrasing of arguments (concerning the harmful effects of smoking) without criticizing in any way, implying that the subject accepts the material although he does not explicitly say so. Example: "So the idea in this statement is that the more a person smokes the bigger the risk he takes."

Every subject was given a score on each of the seven categories, representing the number of paragraphs to which he responded by giving at least one spontaneous association in the given category. In order to check on the reliability of these scores, the protocols were scored independently by one of the authors and by a clerical assistant. The uncorrected reliability coefficients for the seven content analysis categories were found to be .85 or higher.

The only scores used in analyzing these results were those of the clerical assistant, a nonpsychologist who knew nothing about the design of the experiment and who was unfamiliar with any of the hypotheses under investigation. These scores were used rather than pooled scores in order to preclude the possibility of contamination, since the psychologically trained judge could not be prevented from knowing the experimental group to which each subject had been assigned.

The responses to the key interview question dealing with attitude toward smoking were rated on a seven-point scale ranging from *unqualifiedly pro-smoking* (positive comments only, with no signs of conflict or doubt) through *balanced* (negative comments as strong as positive) to *unqualifiedly anti-smoking*. Separate ratings were obtained for each subject's precommunication response and his post-communication response in order to assess the amount of attitude change. These ratings were made independently by the psychologically trained judge and the untrained clerical assistant, under completely blind conditions, i.e., the responses were typed on separate sheets and presented to the judges one at a time in random order so that they did not know whether the protocol being rated was obtained before or after the communication, nor did they know whether the subject was in the Low or High Threat group. A reliability coefficient of .90 was obtained for these ratings. Since there was no possibility of contamination for either judge, the attitude change results were based on the pooled ratings from both judges.

RESULTS AND DISCUSSION

The resistance hypothesis predicts that if a strong threat appeal evokes a relatively high level of emotional tension, it will induce more resistance during exposure to the persuasive message than a mild threat appeal. In order to test this hypothesis, it is necessary first to determine whether the experimental variation of Low versus High Threat elaboration succeeded in inducing differences in affective reactions. Once this is established, we can then proceed to examine the evidence bearing on resistance reactions during exposure to the communication, as well as overall attitude changes.

Affective Arousal

First we shall examine the two indicators of affective reactions obtained from the content analysis of the subjects' spontaneous associations elicited during exposure to the communication. The mean number of paragraphs that evoked expressions of affective disturbance was 1.36 for the High Threat group and 0.12 for the Low Threat group ($t = 2.9$, $p < .01$).[3] Eleven of the 14 subjects (79%) in the High Threat group gave one or more such responses whereas only 2 of the 17 subjects (12%) in the Low Threat group did (corrected chi square = 11.52, $p < .01$). The significantly higher incidence of affective disturbance responses in the High Threat group was attributable to the presence of the seven special paragraphs that said nothing at all about smoking but delineated the distress-

[3] All probability values reported in this study are based on one-tailed tests, since the findings were used to test hypotheses that specified the direction of the differences.

ing symptoms and poor prognosis of cancer in a way that was intended to function as a strong threat appeal. These paragraphs evoked strong expressions of emotional arousal (e.g., "that is shocking," "it's terribly scary").

The second criterion was number of associations containing references to the dangers of cancer. As expected, the mean for the High Threat group was significantly higher than for the Low Threat group (2.86 versus 1.53, $t = 2.02$, $p < .03$). Again, the high incidence of such references occurred primarily in the associations to the seven special paragraphs which elaborated on the distressing things that happen to cancer victims.

Thus, the experimental manipulation appears to have been successful, since the High Threat version of the communication elicited more expressed concern about the dangers of cancer, as well as more expressions of affective disturbance, than the Low Threat version.

Changes in Attitude toward Smoking

Although the purpose of this experiment was to investigate resistances evoked during exposure to the communication, an attempt was made to ascertain whether or not the overall attitude changes induced by the two versions of the communication were in line with the previously reported findings that had led to the formulation of the resistance hypothesis (see pp. 403–404). The blind ratings of favorable and unfavorable comments about smoking in the pre- and postcommunication interviews were used to determine whether each subject was positively or negatively influenced by the communication as a whole. The results indicate that, as expected, the strong threat version tended to be less effective than the mild threat version. Complete interview records were available for 11 subjects in the High Threat group and for 16 subjects in the Low Threat group. (Due to technical recording difficulties, the interview protocols were not available for a few subjects and hence the analysis of attitude change is based on a somewhat smaller number of cases than the other findings.) We found that in the High Threat group, 6 subjects changed in the antismoking direction while 2 changed in the prosmoking direction (Net change = $(6 - 2)/11 \times 100 = 36.4\%$). In contrast, 11 subjects in the Low Threat group changed in the antismoking direction and none changed in the prosmoking direction (Net change $= (11 - 0)/16 \times 100 = 68.8\%$). The 32.4% difference between the net changes yields a *CR* of 1.35 ($p < .09$), according to the formula in Hovland, Lumsdaine, and Sheffield (1949, p. 304). Thus, a difference was found in the expected direction and it approaches significance at slightly beyond the 10% confidence level. This result, when evaluated along with the lower p values reported by

TABLE 1

EFFECTS OF HIGH VERSUS LOW THREAT APPEALS ON EVALUATIVE REACTIONS IN THE SPONTANEOUS ASSOCIATIONS ELICITED DURING EXPOSURE TO AN ANTISMOKING COMMUNICATION[a]

Type of evaluative reaction	Mean number of paragraphs eliciting the given type of evaluative reaction		
	High Threat version ($N = 14$)	Low Threat version ($N = 17$)	t
1. Major criticisms: explicit rejection	3.86	2.24	2.20***
2. Minor criticisms: unfavorable comments about style or objectivity	1.07	1.17	<1.0
3. Major favorable comments: explicit acceptance	6.21	5.36	<1.0
4. Minor favorable comments: praise of style or objectivity	0.93	1.83	1.65*
5. Paraphrasing of arguments	0.29	1.30	1.85**
Total evaluative reactions	12.35	11.90	<1.0

[a] Based on a content analysis of responses to 15 standard paragraphs included in both versions of the communication.
* $p < .06$, one-tailed.
** $p < .05$, one-tailed.
*** $p < .02$, one-tailed.

Janis and Feshbach (1953) and by Haeffner (1956), contributes to the general conclusion that when a strong threat appeal evokes a high degree of affective disturbance, it will tend to be less effective in inducing attitude change than a milder threat appeal.

Critical Reactions

In view of the observed difference in affective arousal, we are able to test the main prediction derived from the resistance hypothesis, namely, that the group exposed to the strong threat appeal will display more critical reactions during exposure to the persuasive communication than the group exposed to the mild appeal. Table 1 shows the pertinent data from the content analysis of evaluative comments in the subjects' spontaneous associations. This analysis was limited to the 15 standard paragraphs included in both versions of the communication, so that there was equal opportunity for both groups to make critical comments. The total number of evaluative comments (shown in the last row of the table) turned out to be nearly equal in the two groups.

The results in the table show that the High Threat group: made significantly more explicit rejection statements ($p < .02$); gave fewer favorable comments about style, etc. ($p < .06$); and made significantly fewer paraphrasing statements concerning the harmfulness of smoking ($p < .05$). Thus, on three of the five evaluative categories, significant or near significant differences were obtained in the predicted direction.[4] For the other two categories, the differences are slight and nonsignificant ($p > .30$).

Although major and minor criticisms are both indicators of negative reactions to the communication, the first category appears to have more face validity for assessing resistance to the communication because it involves clear-cut objections or disagreements. Moreover, a comparison of subjects who changed their attitude after exposure to the communication with those who did not change revealed that the Major Criticisms category was the only evaluative category that yielded a large and statistically significant difference. The other evaluative categories, including Minor Criticisms, yielded small but nonsignificant differences in the expected direction. Hence Major Criticisms was found to be more predictive of overall unresponsiveness to the communication than any of the other types of evaluative response. The finding that a significantly larger number of major criticisms was made by the group exposed to the strong threat version than by the group exposed to the mild threat version supports the hypothesis that a strong threat appeal tends to arouse more resistance than a milder threat appeal.

The fact that the two groups gave an almost equal number of minor criticisms indicates that there was no compensatory difference in the less direct forms of criticisms

[4] A similar pattern of differences would be expected when those who smoke regularly are compared with those who seldom or never smoke, since the communication would tend to arouse greater conflict in the former group: The threat material would be more ego involving for the smokers and might arouse more concern about the harmful consequences of the smoking habit, but their pre-existing attitudes would be more at variance with the recommendation put forth in the communication. On the basis of their precommunication interview reports on the average number of cigarettes smoked each day, the 31 subjects in the two experimental groups were split as close to the median as possible into a subgroup of 14 smokers (more than five cigarettes per day) and 17 nonsmokers (less than five per day). The content analysis data for these two subgroups have been published elsewhere in a theoretical analysis of preconscious affective mechanisms that enter into the resolution of decisional conflicts (Janis, 1959, p. 215). The main finding was that the smokers made significantly fewer statements of explicit acceptance (4.52 versus 6.71, $p < .05$). The smokers also made more statements of explicit rejection than the nonsmokers (3.50 versus 2.53), but the difference was not large enough to be statistically significant. (In the previously published table [p. 215] there are two typographical errors in the bottom row, which shows the means for explicit rejection; the significance test data and all other means shown in the table are correct.) In general, the findings bear out the expectation that smokers will tend to be more resistant than nonsmokers to a communication which calls attention to the dangers of smoking. The observed differences between smokers and nonsmokers, however, could not account for the observed differences between the High Threat and Low Threat groups inasmuch as the subjects in the two groups were very closely equated on the basis of their precommunication statements about the average number of cigarettes they smoked each day (see p. 407 above).

that could counterbalance the predominantly higher incidence of major criticisms found in the group exposed to the strong threat version. It is also noteworthy that the latter group gave significantly fewer paraphrases of the arguments than the group exposed to the milder threat version. This finding suggests that the strong threat appeal induced an inhibitory tendency such that the subjects avoided thinking over the assertions in the communication concerning the unfavorable consequences of smoking. Thus, the results concerning paraphrasing of arguments, like those bearing on major criticisms, can be readily interpreted in terms of the following general proposition: The more strongly fear is aroused by a warning communication, the more strongly motivated the person will become to avoid symbolic responses and thought sequences which lead him to recall or to focus his attention on the essential content of the arguments and conclusions. The evidence in Table 1 suggests that the avoidance tendency was somewhat stronger in the High Threat group than in the Low Threat group; but this tendency does not appear to have been dominant during exposure to all paragraphs of the communication, especially since there was an equally high incidence of explicit acceptance in both groups.

Major and minor favorable comments (Rows 3 and 4 in the table), which are presumed to be indicators of acceptance reactions that facilitate attitude change, occurred with a high frequency in both experimental groups. In fact, for both versions of the communication, the combined total of favorable comments exceeds the combined total of unfavorable comments. It must be emphasized, therefore, that the insertion of the strong threat appeal, while reducing the overall influence of the communication on the audience, did not make the communication totally ineffective. We have already noted in the attitude change data (p. 407) that the High Threat version succeeded in influencing a minority of the audience but the net percentage who changed was significantly smaller than in the audience exposed to the Low Threat version. Thus, it cannot be said that the heightened resistance created by adding the strong threat appeal completely outweighed the effectiveness of the arguments contained in the standard portions of the communication; rather, the strong threat material had a dampening effect, preventing the rest of the communication from being as strongly effective as it otherwise would have been.[5]

SUMMARY AND CONCLUSIONS

An experiment was designed to test the implications of earlier studies by Janis and Feshbach (1953, 1954) bearing on resistance to attitude change induced by the use of fear arousing appeals in persuasive communications. The present study was carried out with 31 adults, age 18–55, representing a broad range of occupations. On a stratified random basis, the subjects were assigned to two experimental groups, both of which were given an antismoking communication which promoted the conclusion that heavy smoking causes cancer and which recommended that everyone should avoid or cut down on smoking. One group ($N = 17$) was given a Low Threat version, consisting of 15 paragraphs devoted mainly to authentic quotations from medical authorities. The other group ($N = 14$) was given a High Threat version, consisting of the same 15 paragraphs supplemented by 7 special "threat" paragraphs that gave vivid descriptions of the pain, suffering, and poor prognosis of cancer victims.

The communications were presented in a private session during which the subject was asked to express his thoughts and feelings after reading each paragraph. In order to eliminate cues that might give a distorted picture of the subject's cognitive and emotional reactions, a special device (auditory

[5] It seems probable that the heightened resistance evoked by the strong threat version involves a preconscious or unconscious defensive process. But the question of how aware the subjects were of the purpose of their resistant reactions cannot be answered from the evidence currently at hand. It should be mentioned that none of the subjects ever spontaneously asserted that he was deliberately trying to avoid being influenced by the fear arousing contents, nor did any subject admit to having had such an intention, when interviewed at the end of the session. The absence of such admissions, however, could be the result of deliberate withholding of information and hence does not necessarily imply a *lack of conscious intention* to ward off the distressing emotional impact of the fear arousing communication.

feedback suppressor) was used which enabled the subject to give his spontaneous associations aloud without hearing the sound of his own voice. Tape recordings of each subject's verbalizations were analyzed according to a systematic content analysis procedure. These results were supplemented by an analysis of changes in the subjects' responses in intensive interviews conducted before the communication and again afterwards.

The main conclusions were as follows:

1. The strong threat version of the communication elicited more verbal expressions of emotional tension and more concern about the danger of cancer than the mild appeal, thus indicating that the intended experimental manipulation was successful.

2. In line with earlier experiments in which mild threat appeals were found to produce more attitude change than strong threat appeals, the mild threat version of the antismoking communication tended to be more effective than the strong threat version, as assessed by blind ratings of the attitudes expressed by subjects in pre- and postcommunication interviews ($p < .10$).

3. During the period of exposure to the communication, there were more manifestations of resistance in the High Threat group than in the Low Threat group. The main criterion for assessing resistance was the number of explicit rejection statements. The High Threat group made many more such statements than the Low Threat group ($p < .02$). The former group also made fewer favorable comments about the style or objectivity of the communication ($p < .06$) and gave fewer paraphrasing responses of the type which indicate implicit acceptance ($p < .05$). No significant differences were obtained on number of explicit acceptance statements and number of minor criticisms concerning the style or objectivity of the communication.

The findings were interpreted as supporting the following general hypothesis: *When a relatively high level of fear is induced by the warnings presented in a persuasive communication, the recipients will become motivated to develop psychological resistances to the communicator's arguments, conclusions, and recommendations.* This hypothesis had been inferred from earlier experimental findings based solely on measures of communication effectiveness obtained by comparing pre- and postcommunication questionnaire responses. The present study provides more direct measures which indicate that resistances tend to be mobilized by a strong threat appeal during the period of exposure to the communication.

REFERENCES

GOLDSTEIN, M. The relationship between coping and avoiding behavior and attitude modification. Unpublished doctoral dissertation, University of Washington, 1957.

GOLDSTEIN, M. The relationship between coping and avoiding behavior and response to fear-arousing propaganda. *J. abnorm. soc. Psychol.*, 1959, **58**, 247–252.

HAEFFNER, D. Some effects of guilt-arousing and fear-arousing persuasive communications on opinion change. Unpublished Technical Report, August 15, 1956, Office of Naval Research, Contract Number N6 onr-241. (Abridgement of unpublished doctoral dissertation, University of Rochester, 1956)

HOVLAND, C. I., LUMSDAINE, A. A., & SHEFFIELD, F. D. *Experiments on mass communication*. Princeton, N. J.: Princeton Univer. Press, 1949.

JANIS, I. Motivational factors in the resolution of decisional conflicts. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1959*. Lincoln: Univer. Nebraska Press, 1959. Pp. 198–231.

JANIS, I., & FESHBACH, S. Effects of fear-arousing communications. *J. abnorm. soc. Psychol.*, 1953, **48**, 78–92.

JANIS, I., & FESHBACH, S. Personality differences associated with responsiveness to fear-arousing communications. *J. Pers.*, 1954, **23**, 154–166.

JANIS, I., & MILHOLLAND, H. The influence of threat appeals on selective learning of the content of a persuasive communication. *J. Psychol.*, 1954, **37**, 75–80.

MOLTZ, H., & THISTLETHWAITE, D. Attitude modification and anxiety reduction. *J. abnorm. soc. Psychol.*, 1955, **50**, 231–237.

(Received December 19, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 411–418

PHYSIOLOGICAL AND SUBJECTIVE RESPONSES TO PAIN PRODUCING STIMULATION UNDER HYPNOTICALLY-SUGGESTED AND WAKING-IMAGINED "ANALGESIA" [1]

THEODORE XENOPHON BARBER AND KARL W. HAHN, JR.

State Hospital, Medfield, Massachusetts *School of Medicine, Boston University*

With a few possible exceptions, experimental studies (Brown & Vogel, 1938; Doupe, Miller, & Keller, 1939; Dynes, 1932; Levine, 1930; Sears, 1932; Shor, 1959; Sutcliffe, 1961; West, Niell, & Hardy, 1952) designed to test the hypothesis that "hypnotically-suggested analgesia" is effective in attenuating pain have been limited in scope as follows:

1. The pain producing stimuli employed (pinprick, electric shock, or radiant heat applied to a limb for no more than 3 seconds) were of brief or momentary duration.

2. Pain reactivity under hypnotically-suggested analgesia was compared with reactivity under an uninstructed condition. Although in general it was found that suggestions of analgesia given under hypnosis were more effective than no instructions in reducing subjective and physiological responses to noxious stimulation of short duration, this does not demonstrate that hypnosis was a necessary condition in producing this effect. As Brown and Vogel have suggested, it may be possible to produce a similar reduction in response to pain by instructing a control group to try to imagine that a noxious stimulus is nonpainful (cf. Barber, 1959b, 1960).

The present experiment was designed to determine if hypnotically-suggested analgesia is more effective than "waking-imagined analgesia" in attenuating subjective and physiological responses to a stimulus which elicits pain over a longer period of time than the stimuli used in earlier studies. Water near the freezing point applied to a limb for 3 minutes was selected as the pain stimulus. Previous investigators (Benjamin, 1958; Engel, 1956; Kunkle, 1949; McMurray, 1950; Wolf & Hardy, 1941) had reported that "aching pain" is elicited in normal subjects by this stimulus within 10–60 seconds; if the stimulus is not removed, pain continues for 2–4 minutes before adaptation sets in; and the intensity of the pain experienced is closely related to increments in heart rate, systolic and diastolic pressure, and respiratory irregularities.

[1] This research was supported by a grant (MY4825) from the National Institute of Mental Health, United States Public Health Service.

We are indebted to Joan Burke and David S. Calverley for pretesting the subjects on the eight-point Suggestibility scale.

METHOD

Selection of Subjects

Forty-eight subjects, homogeneous with respect to age, sex, and "suggestibility," were selected as follows. A Suggestibility scale consisting of eight standardized test suggestions (e.g., arm heaviness, inability to unclasp hands, body immobility) was administered (by experimenters not participating in the present study) to 192 women college students and nursing students under nonhypnotic conditions, as described by Barber and Glass (1962, Study I). The 48 subjects (25%) scoring 5.5 or higher on this eight-point scale, and thus falling in the upper quartile of the pretested population with respect to "suggestibility," were offered $3.00 to participate in the present experiment which was defined as a "physiological study." All subjects agreed to participate. The age range was 17–23 with a median of 19. The 48 subjects were allocated at random to four experimental treatments (Hypnosis, Waking-Imagination, Uninstructed, and Control) with 12 subjects to each treatment.

Procedure

The experiment was conducted in a laboratory maintained at $23 \pm 1°$ centigrade. The preliminary period (15 minutes) consisted of adjustment of polygraph channels and application of leads for recording the following: *muscle action potentials* from electrodes fixed by a head band and placed over the *frontalis* muscle approximately 3 centimeters above the eyebrows; *respiration* by chest pneumograph actuating a Grass PT5 volumetric pressure transducer; *heart rate* by electrocardiograph lead I with electrodes at the right wrist and immediately below the left elbow; and *skin resistance*

from silver chloride electrodes at the right palm. Upon completion of these preliminaries, the following treatments were imposed while the physiological measures were continuously and simultaneously recorded on Grass polygraph equipment.

Hypnotically-suggested analgesia. Subjects allocated to the hypnosis treatment were given a standardized 20-minute hypnotic induction consisting of eye fixation on a light blinking in synchrony with the sound of a metronome, repetitive suggestions of eye closure, relaxation, drowsiness, and sleep, and suggestions designed to produce positive motivation to perform maximally on forthcoming tasks. To assess the effectiveness of the hypnotic induction procedure, the eight-point standardized Suggestibility scale which had been used in the pretest was then administered. The subjects appeared to enter trance (i.e., appeared drowsy and showed psychomotor retardation and a lack of spontaneity and initiative) and responded positively to the test suggestions (passing test suggestions which they had failed in the pretest). Suggestions were then given for a period of 1 minute to induce insensitivity in the left hand, as follows:

> As you listen to my voice you begin to notice a lack of feeling in your left hand. You notice that your left hand is becoming numb and insensitive. As I speak to you your left hand becomes more and more numb; you can feel less and less with it; it is losing all feeling, all sensation. Very soon now you will not be able to feel anything in the hand, it will be without sensation. I will count to three and at the count of three the hand will be completely numb, insensible, and without feeling. One. The hand is losing all feeling, all sensation, it is becoming more and more numb, you feel less and less with it, you are losing awareness of it, soon you will be unable to feel anything at all with it. Two. The hand is quite numb now, there is hardly any feeling in it, it is as if it were asleep. Three. The hand is now numb, insensitive, and without feeling; it cannot feel anything, it is without sensation. The hand will remain numb and insensitive indefinitely; no matter what I ask you to do the hand will remain numb and insensitive; it will continue to be numb and without feeling until such time as I restore sensitivity to the hand.

After a 60-second pause, the hypnotized subject was told to immerse the insensitive hand up to the wrist in "cold water" (water at 2° centigrade) and not to remove the hand until told to do so. After 3 minutes of immersion the subject was instructed to remove the hand. Suggestions were then given to restore sensitivity to the hand and soon afterwards the subject was awakened.

Nonhypnotic conditions. The three nonhypnotic conditions differed from the above condition in that the hypnotic induction procedure, the Suggestibility scale, and the hypnotic suggestions of insensitivity were not administered. To control the amount of time that subjects remained in the experimental setting, the three nonhypnotic groups were first asked to sit quietly for 20 minutes while the physiological measures were recorded; to substitute for the Suggestibility scale, these groups were asked voluntarily to perform motions similar to those included in the scale, e.g., as a substitute for the test suggestion that the subject's raised arm was becoming heavy and moving down involuntarily, the subject was asked voluntarily to raise and then to lower an arm. Upon completion of these preliminaries, the experimental treatments were as follows:

Waking-imagined analgesia. Subjects allocated to this condition were given the following instructions for a 1-minute period designed to produce positive motivation to imagine a pleasant situation when the noxious stimulus was applied:

> In the next part of the experiment I want you to place your left hand in cold water. If you think about how cold the water is you may find it uncomfortable. However, if you try to imagine that the water is pleasant and try to think of it as *not* uncomfortable, you will be able to keep your hand in the water without being bothered by it at all. Everyone who tried to imagine that the water was pleasant and refused to think of it as uncomfortable was able to do well on this test. You see this test all depends on your willingness to imagine and to try to think in a certain way. If you do not try to the best of your ability to carry out the instructions you will fail the test and ruin this part of the experiment. What I ask then is your cooperation in helping this experiment by carrying out the following instructions: When your hand is in the water, try to imagine that it is a very hot day, that the water feels pleasantly cool, and that your hand is relaxed and comfortable. I'm sure that you'll be able to continue imagining this and that you will not fail this test.[2]

After a 60-second pause, the subject was told to immerse the left hand in the "cold water" (at 2° centigrade) and not to remove the hand until told to do so; after 3 minutes of immersion the subject was instructed to remove the hand.

Uninstructed condition. Following a 60-second null period (no conversation between subject and experimenter), subjects assigned to this treatment were told to immerse the left hand in "cold water" (at 2° centigrade) and not to remove the hand until

[2] It should be noted that instructions for both the Hypnosis and the Waking-Imagination groups were designed to produce "analgesia" (i.e., to reduce pain responsiveness) indirectly without mentioning the word *pain*. Pilot work appeared to indicate that such indirect instructions are more effective in reducing pain responsiveness than direct suggestions of "analgesia" (e.g., "Your hand is insensitive *to pain*") which at times tend to define the situation to the subject as one in which pain is to be expected.

given further instructions; after 3 minutes, the subjects were instructed to remove the hand.

Control condition. This condition was the same as the Uninstructed condition with one change: although the subjects were told to immerse the left hand in "cold water," the water was actually *at room temperature*.

Assessment of Subjective Experiences

Soon after the stimulus period, all subjects were asked to write the answer to one question:

> Your hand was in the water for three minutes [or, for two minutes, if the subject withdrew the hand before the third minute]. Please indicate how your hand felt while it was in the water by placing a letter [*A* for the first minute your hand was in the water; *B* for the second minute; and *C* for the third minute] in front of the appropriate description:
>
> My hand felt: ——cool; ——cold; ——numb; ——uncomfortable; ——painful; ——very painful.

The subject was given a Subjective score for each estimated stimulus minute (as indicated by the letters *A*, *B*, or *C*) as follows: if "very painful" was checked the score was 4; if "very painful" was not checked but "painful" was checked, the score was 3; if neither "very painful" nor "painful" were checked but "uncomfortable" was checked, the score was 2; and if neither "very painful," "painful," nor "uncomfortable" were checked, the score was 1. (If the hand was withdrawn from the water before the 3-minute stimulus period was completed, the subject received a score of 4 for each stimulus minute in which the hand was not in the water.) Each subject thus received three Subjective scores on this question, one for each estimated stimulus minute.

Criteria of pain responsiveness consisted of: period of time the subject allowed the hand to remain in the water before withdrawal; subjective reports of pain experienced; and alterations in physiological functions (muscle tension, respiration, heart rate, and skin resistance).

RESULTS

Time-for-Withdrawal

Two subjects in the Hypnosis group, one in the Waking-Imagination group, and two in the Uninstructed group withdrew the hand (from the water at 2° centigrade) before the 3-minute stimulus period was completed. All subjects in the Control group kept their hand in the water at room temperature for the required 3 minutes. The mean number of seconds before withdrawal was 167, 168, 160, and 180 under the Hypnosis, Waking-Imagination, Uninstructed, and Control conditions, respectively. Since the time-for-withdrawal scores did not meet the assumptions for a parametic analysis (homogeneity of variances and normality) the Kruskal-Wallis analysis of variance by ranks for independent samples (Siegel, 1956) was used to determine the significance of the differences among the four groups. This test indicated that the groups did not differ significantly on this measure ($H = .70$, $p > .80$).

Subjective Reports

The mean Subjective scores for each estimated stimulus minute are presented in Table 1. As this table indicates the Hypnosis and Waking-Imagination groups generally showed Subjective scores around 2.0 ("uncomfortable" level), the Uninstructed group showed Subjective scores around 3.0 ("painful" level), and the Control group showed Subjective scores around 1.0 (below the "uncomfortable" level).

Since the Subjective scores were neither normally distributed nor homogeneous with respect to variances, they were analyzed by Kruskal-Wallis tests and by Mann-Whitney *U* tests (Siegel, 1956). The results of the Kruskal-Wallis tests, summarized in Table 1, indicated that the mean Subjective scores attained by the four groups differed significantly beyond the .001 level. The results of the Mann-Whitney *U* tests, also summarized in Table 1, indicated the following:

TABLE 1

SUBJECTIVE SCORES UNDER HYPNOTICALLY-SUGGESTED ANALGESIA, WAKING-IMAGINED ANALGESIA, AN UNINSTRUCTED CONDITION, AND A CONTROL CONDITION

Estimated stimulus minute	Experimental conditions				H[a]
	Hypnotically-suggested analgesia	Waking-imagined analgesia	Uninstructed	Control	
First minute	1.6_{ab}	1.8_{b}	2.8_{c}	1.1_{a}	21.47[b]
Second minute	2.2_{b}	1.8_{ab}	3.1_{c}	1.1_{a}	23.64[b]
Third minute	2.2_{b}	2.4_{b}	3.0_{b}	1.0_{a}	16.70[b]

Note.—Means in the same row containing the same subscript letter are not significantly different from one another at the .05 level by the Mann-Whitney *U* test. Any two row means which do not contain a common subscript differ significantly from one another at or beyond the .05 level by the Mann-Whitney *U* test.

[a] Kruskal-Wallis test.

[b] $p < .001$.

1. The Hypnosis and Waking-Imagination groups did not differ significantly in reported pain experience.

2. With the exception of the third estimated stimulus minute, the Hypnosis and Waking-Imagination groups reported significantly less discomfort and pain than the Uninstructed group.

3. During two of the three estimated stimulus minutes, the Hypnosis and Waking-Imagination groups reported significantly more discomfort and pain than the Control group.

4. The Uninstructed group reported significantly more discomfort and pain than the Control group during all three estimated stimulus minutes.

TABLE 2

AUTONOMIC LABILITY SCORES IN RESPONSE TO NOXIOUS STIMULATION UNDER HYPNOTICALLY-SUGGESTED ANALGESIA, WAKING-IMAGINED ANALGESIA, AN UNINSTRUCTED CONDITION, AND A CONTROL CONDITION

Physiological measures	Experimental conditions				F
	Hypnotically-suggested analgesia	Waking-imagined analgesia	Uninstructed	Control	
Forehead muscle potentials	48_b	47_b	58_a	47_b	3.66*
Respiratory irregularity	51_b	51_b	56_a	42_c	4.28**
Heart rate	53_a	53_a	52_a	42_b	3.98*
Skin resistance	48_b	47_b	48_b	57_a	4.77**

Note.—Means in the same row containing the same subscript letter are not significantly different from one another at the .05 level by the Duncan test. Any two row means which do not contain a common subscript differ significantly from one another at or beyond the .05 level by the Duncan test.

* $p < .05$.

** $p < .01$.

Physiological Responses

The following physiological measures were each computed four times from the subject's polygraph chart, once for the 60-second prestimulus period (base levels) and once for each of the three stimulus minutes (response levels): forehead muscle tension as measured by time in seconds that the integrated electromyographic tracing was more than double the subject's base level of activity; respiratory irregularity in terms of percentage of respiratory cycles deviating by ± 25% in rate or amplitude from the subject's resting respiratory pattern as manifested during the preceding 20-minute quiescent period; heart rate in beats per minute; and highest level of skin resistance in ohms. The maximum score attained on each physiological measure during any one of the three stimulus minutes (or during any one of the two stimulus minutes if the hand was withdrawn from the stimulus before the third minute) was selected as the response level for analysis. To remove the effect of prestimulus (base) levels on maximal response levels, the data were converted to Autonomic Lability scores with a mean of 50 and a standard deviation of 10 (Lacey, 1956). Since the Autonomic Lability scores met the restrictive assumptions for a parametric analysis, they were treated by analyses of variance, and Duncan (1955) multiple range tests were used to compare the means. The F ratios from the variance analyses, as summarized in Table 2, showed that all four sets of Autonomic Lability scores differed significantly at or beyond the .05 level. The results of the Duncan tests, summarized in Table 2, indicated the following:

1. The Hypnosis and Waking-Imagination groups did not differ significantly on any physiological response to the noxious stimulus.

2. As compared to the Uninstructed condition, hypnotically-suggested analgesia and waking-imagined analgesia were effective in reducing muscle tension and respiratory irregularities.

3. Hypnotically-suggested analgesia and waking-imagined analgesia were effective in reducing muscle tension but not respiratory irregularities to the low level found under the Control condition.

4. The Hypnosis and Waking-Imagination conditions did not differ from the Uninstructed condition in heart rate and skin resistance levels; under these conditions subjects showed significantly faster heart rate and significantly lower skin resistance than under the Control condition.

Relation of Physiological Response to Subjective Reports

Product-moment coefficients of correlation were computed between the maximum Subjective scores (highest Subjective score attained by the subject during any one of the

three estimated stimulus minutes) and each of the four sets of Autonomic Lability scores. A significant negative correlation ($r = -.46$) was obtained between the maximum Subjective scores and the skin resistance Autonomic Lability scores, and significant positive correlations were obtained between the maximum Subjective scores and the muscle tension, respiratory irregularity, and heart rate Autonomic Lability scores ($r = .29$, .44, and .29, respectively). The obtained correlations thus appeared to be in line with the results of previous studies, summarized by Woodworth and Schlosberg (1954) and Beecher (1959), which suggest that the experience of pain is associated with a decrease in skin resistance and an increase in muscle tension, respiratory irregularities, and heart rate.

DISCUSSION

In designing this experiment one major assumption was made and one question was asked. The assumption was that the stimulus selected for study (water near the freezing point applied to a limb for 3 minutes) would produce pain in an uninstructed group. The question asked was: Is hypnotically-suggested analgesia more effective than waking-imagined analgesia in attenuating the pain produced by this stimulus? The experimental findings will be discussed with respect to the background assumption and the question of interest. The findings will also be related to other recent research concerned with hypnosis.

Effectiveness of Water at 2° Centigrade as a Pain Stimulus

A comparison of the Subjective scores and the Autonomic Lability scores of the Uninstructed group (which received the 2° centigrade stimulus) and the Control group (which received stimulation at room temperature) supports the assumption that water near the freezing point is an effective pain producing stimulus. The Uninstructed group reported that the stimulus was experienced as painful. The Uninstructed group also showed physiological reactions during the stimulation (elevations in muscle tension, heart rate, and respiratory irregularities and reduction in skin resistance) which are generally found in close association with painful stimulation and were not shown by the Control group.

Pain Reactivity as Affected by Hypnotically-Suggested Analgesia and Waking-Imagined Analgesia

As compared to the Uninstructed condition, hypnotically-suggested analgesia was effective in eliciting reports of reduced pain experience and in attenuating some but not all physiological responses to the stimulus. It should be noted that:

1. Although the Hypnosis group reported less discomfort and pain than the Uninstructed group, discomfort and pain were not abolished; the Hypnosis group reported significantly more discomfort and pain than the Control group which received nonnoxious stimulation.

2. Although hypnotically-suggested analgesia was effective in reducing muscle tension and respiratory irregularities, it had no measurable effect on the skin resistance and cardiac acceleratory responses to the stimulus.

3. The responses that were reduced under hypnotically-suggested analgesia (subjective reports, forehead muscle tension, and respiratory irregularities) appear to be normally amenable to some degree of voluntary control; the responses that were not affected (cardiac acceleration and reduction in skin resistance) are more closely associated with activity of the autonomic nervous system and appear to be less amenable to voluntary control.

The major finding from this experiment is that hypnotically-suggested analgesia appeared to be no more effective than waking-imagined analgesia in reducing pain reactivity. Both conditions produced reports of reduced pain experience; both conditions were associated with a reduction in muscle tension and respiratory irregularities; and both conditions were ineffective in altering two autonomic responses (cardiac acceleration and drop in skin resistance) which are normally elicited by water near the freezing point. The data thus appear to indicate that pain reactivity may be as effectively reduced either by administering suggestions of analgesia under a hypnotic condition or by attempting to motivate the subject under a nonhypnotic

condition to imagine a pleasant situation when noxious stimulation is applied.

Relation of Findings to Other Recent Research in Hypnosis

A series of recent experiments, summarized below, which compared response to suggestions under an hypnotic condition and under one or more nonhypnotic conditions, obtained results which are in line with the results of the present experiment.

Barber and Calverley (1962) compared responses to eight standardized test suggestions—e.g., arm heaviness, inability to unclasp hands, thirst "hallucination," body immobility—under an Hypnosis condition, a Motivated condition, an Uninstructed condition, and three other control conditions. Subjects allocated to the Hypnosis condition were assessed on the test suggestions after a standardized 15-minute hypnotic induction procedure had been administered and when the subjects appeared to be in trance. Subjects assigned to the Motivated condition were assessed on the test suggestions after they were told that: they were to be tested for ability to imagine; others had performed well on the imaginative tasks; they also could perform well if they tried to imagine vividly; if they did not try to the best of their ability to imagine, they would fail the test and the experiment would be worthless. In harmony with previous findings reported by Weitzenhoffer and Sjoberg (1961) and Barber and Glass (1962), the Hypnosis group was more "suggestible" (i.e., attained higher scores on the test suggestions) than the Uninstructed group and the other three control groups; however, the Hypnosis group was *not* more responsive to the test suggestions than the group that had been given brief instructions designed to produce positive motivation to perform well on the "imaginative" tasks. The essential features of this study were replicated in a cross-validation experiment (Barber & Calverley, 1962) and confirmatory findings were obtained. A standardized 15-minute hypnotic induction procedure appeared to be no more effective than brief motivating instructions in producing high levels of overt response to the eight standardized test suggestions and in eliciting subjective reports that the suggested effects were experienced.

In an additional experiment Barber and Calverley (1962) obtained data which strongly suggest that overt behavior indicative of "amnesia" and subjective reports that experimental events were "actually forgotten" can be as effectively elicited by attempting to motivate nonhypnotic subjects to try to forget the experimental events as by suggesting to hypnotized subjects (i.e., to subjects who have received an hypnotic induction and who appear to be in trance) that they shall forget the experimental events. Along similar lines, Barber (1959a) presented data indicating that a performance that can at times be evoked by appropriate suggestions given to hypnotized subjects—"hallucinating" colors on a gray card, and then reporting complementary colored afterimages of the hallucinated colors (Erickson & Erickson, 1938; Rosenthal & Mele, 1952)—can also be evoked from some control subjects by instructions to "try very hard to imagine the (specified) colors." Barber and Deeley (1961) found that a control group instructed to "try as hard as you possibly can not to see the color red (or the color green) on the Ishihara plates" performed on the Ishihara test in a similar manner as "hypnotic color-blind" subjects. London and Fuhrer (1961) demonstrated that enhanced performance on a hand dynamometer and a weight endurance task could be elicited as effectively either by suggesting increased strength and endurance to hypnotized subjects or by motivating a control group to perform maximally on the tasks.

In brief, the results of the present experiment and the findings obtained in other recent experiments appear to indicate the following: If the term hypnosis refers to the administration of one of the procedures that have been historically categorized as hypnotic inductions and to the appearance of trance on the part of the subject, then hypnosis may be no more effective than a motivated nonhypnotic condition in producing positive responses to suggestions of analgesia, arm heaviness, inability to unclasp hands, thirst hallucination, body immobility, amnesia, color blindness, "color hallucinations," and

enhanced strength and endurance.[3] Additional experiments are needed to determine if subjects who have received an hypnotic induction and who show signs of trance are more responsive to other types of suggestions (e.g., age regression, negative visual or auditory hallucinations) than control subjects who are motivated to perform well on the suggested tasks (cf. Barber, 1961, 1962a, 1962b, in press-a, in press-b).

SUMMARY

Waking-imagined analgesia was found to be as effective as hypnotically-suggested analgesia in attenuating pain (produced by immersion of a limb in water near the freezing point for 3 minutes) as indicated by subjective reports of reduced pain experience and by reduction in respiratory irregularities and forehead muscle tension. Neither waking-imagined analgesia nor hypnotically-suggested analgesia affected two autonomic responses (cardiac acceleration and reduction in skin resistance) which are normally elicited by painful stimulation.

[3] At times the word "hypnosis" is used in a much broader sense to refer to heightened response to suggestions, irrespective of how this heightened response is brought about. If the word hypnosis is to be used in this way, the results of the present experiment and of other recent experiments can be interpreted as indicating that an "hypnotic induction procedure" and the appearance of "trance" on the part of the subject are not necessary for hypnosis. However, it appears to us that using the word hypnosis to refer to heightened response to suggestions, irrespective of how this heightened response is brought about, results in serious confusion because of the close historical association of the word hypnosis with hypnotic induction procedures and with the signs of trance. Research concerned with the determinants of "suggestibility" could avoid ambiguity if the term hypnosis is used to refer to the administration of one of the procedures that have been traditionally categorized as hypnotic inductions and to the appearance of trance (e.g., lethargy, lack of spontaneity) on the part of the subject, and if the term hypnosis is *not* applied to other experimental conditions in which heightened suggestibility may be observed, such as to the above experimental conditions in which subjects were given brief motivating instructions.

REFERENCES

BARBER, T. X. The afterimages of "hallucinated" and "imagined" colors. *J. abnorm. soc. Psychol.*, 1959, **59**, 136–139. (a)

BARBER, T. X. Toward a theory of pain: Relief of chronic pain by prefrontal leucotomy, opiates, placebos, and hypnosis. *Psychol. Bull.*, 1959, **56**, 430–460. (b)

BARBER, T. X. "Hypnosis," analgesia, and the placebo effect. *J. Amer. Med. Ass.*, 1960, **172**, 680–683.

BARBER, T. X. Physiological effects of "hypnosis." *Psychol. Bull.*, 1961, **58**, 390–419.

BARBER, T. X. Experimental controls and the phenomena of "hypnosis": A critique of hypnotic research methodology. *J. nerv. ment. Dis.*, 1962, **134**, 493–505. (a)

BARBER, T. X. Hypnotic age-regression: A critical review. *Psychosom. Med.*, 1962, **24**, 286–299. (b)

BARBER, T. X. Toward a theory of hypnosis: Posthypnotic behavior. *Arch. gen. Psychiat.*, in press. (a)

BARBER, T. X. Toward a theory of "hypnotic" behavior: The "hypnotically-induced dream." *J. nerv. ment. Dis.*, in press. (b)

BARBER, T. X., & CALVERLEY, D. S. "Hypnotic behavior" as a function of task motivation. *J. Psychol.*, 1962, **54**, 363–389.

BARBER, T. X., & DEELEY, D. C. Experimental evidence for a theory of hypnotic behavior: I. "Hypnotic color-blindness" without "hypnosis." *Int. J. clin. exp. Hypn.*, 1961, **9**, 79–86.

BARBER, T. X., & GLASS, L. B. Significant factors in hypnotic behavior. *J. abnorm. soc. Psychol.*, 1962, **64**, 222–228.

BEECHER, H. K. *Measurement of subjective responses.* New York: Oxford Univer. Press, 1959.

BENJAMIN, F. B. Effect of aspirin on suprathreshold pain in man. *Science*, 1958, **128**, 303–304.

BROWN, R. R., & VOGEL, V. H. Psychophysiological reactions following painful stimuli under hypnotic analgesia contrasted with gas anaesthesia and novocain block. *J. appl. Psychol.*, 1938, **22**, 408–420.

DOUPE, J., MILLER, W. R., & KELLER, W. K. Vasomotor reactions in the hypnotic state. *J. Neurol. Psychiat.*, 1939, **2**, 97–102.

DUNCAN, D. B. Multiple range and multiple *F* tests. *Biometrics*, 1955, **11**, 1–42.

DYNES, J. B. An experimental study of hypnotic anesthesia. *J. abnorm. soc. Psychol.*, 1932, **27**, 79–88.

ENGEL, B. T. Physiological correlates of pain and hunger. Unpublished doctoral dissertation, University of California, Los Angeles, 1956.

ERICKSON, M. H., & ERICKSON, ELIZABETH M. The hypnotic induction of hallucinatory color vision followed by pseudo-images. *J. exp. Psychol.*, 1938, **22**, 581–588.

KUNKLE, E. C. Phasic pains induced by cold. *J. appl. Physiol.*, 1949, **1**, 811–824.

LACEY, J. I. The evaluation of autonomic responses: Toward a general solution. *Ann. N. Y. Acad. Sci.*, 1956, **67**, 123–164.

LEVINE, M. Psychogalvanic reaction to painful stimuli in hypnotic and hysterical anesthesia. *Bull. Johns Hopkins Hosp.*, 1930, **46**, 331–339.

LONDON, P., & FUHRER, M. Hypnosis, motivation, and performance. *J. Pers.*, 1961, **29**, 321–333.

MCMURRAY, G. A. Experimental study of a case of insensitivity to pain. *Arch. Neurol. Psychiat.*, 1950, **64**, 650–677.

ROSENTHAL, B. G., & MELE, H. The validity of hypnotically induced color hallucinations. *J. abnorm. soc. Psychol.*, 1952, **47**, 700–704.

SEARS, R. R. Experimental study of hypnotic anesthesia. *J. exp. Psychol.*, 1932, **15**, 1–22.

SHOR, R. Explorations in hypnosis: A theoretical and experimental study. Unpublished doctoral dissertation, Brandeis University, 1959.

SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

SUTCLIFFE, J. P. "Credulous" and "skeptical" views of hypnotic phenomena: Experiments on esthesia, hallucination, and delusion. *J. abnorm. soc. Psychol.*, 1961, **62**, 189–200.

WEITZENHOFFER, A. M., & SJOBERG, B. M., JR. Suggestibility with and without "induction of hypnosis." *J. nerv. ment. Dis.*, 1961, **132**, 204–220.

WEST, L. J., NIELL, K. C., & HARDY, J. D. Effects of hypnotic suggestion on pain perception and galvanic skin response. *Arch. Neurol. Psychiat.*, 1952, **68**, 549–560.

WOLF, S., & HARDY, J. D. Studies on pain: Observations on pain due to local cooling and on factors involved in the "cold pressor" response. *J. clin. Invest.*, 1941, **20**, 521–533.

WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology*. (Rev. ed.) New York: Holt, 1954.

(Early publication received July 11, 1962)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 419–425

CASE REPORT

UNDERSTANDING LANGUAGE WITHOUT ABILITY TO SPEAK: A CASE REPORT[1]

ERIC H. LENNEBERG

Harvard University

Infants' random babbling is generally considered to play a major—by some, an essential—role in the acquisition and development of language. In fact, many psychologists believe that the main reason for the failure of mammals other than man to learn to speak or even to bring their vocalizations under new and varied stimulus control is due to their scanty random vocalizations. Recently it was discovered that dolphins make a great variety of vocal tract produced noises and, to be sure, the hope was soon expressed that these animals may with proper training learn to converse with their trainers in English. Incidentally, this hope was shared by Dr. Doolittle, Hugh Lofting's charming literary creation.

Our understanding of human behavior is often greatly enlightened by careful investigations of clinical aberrations and in many instances disease or congenital abnormalities provide conditions that may replace the crucial experiments on children that our superego forbids us to plan and perform. No psychological theory on verbal behavior can be successful unless it takes into consideration and accounts for the pathological variations that may be observed clinically. In the present paper I will present the case of a child who is typical of a large group of children with deficits in their motor execution of language skills but who can learn to understand language even in the total absence of articulation. This and similar clinical material forces us to review our theoretical formulations concerning the role of babbling and echoic responses in the acquisition of language and to review once more the relationship between understanding and speaking a language.

Case Report

This is an 8-year-old boy who has a congenital disability for the acquisition of motor speech skills (anarthria) which, however, has not impaired his ability to learn to understand language.

Medical History

The mother, at the time of subject's birth, was a para I, gravida II white housewife who had enjoyed good health. Throughout her second pregnancy she suffered from a mild chronic bronchitis but gestation was otherwise unremarkable. The baby was born 2 weeks prematurely by dates. Delivery was spontaneous, preceded by 6 hours of labor though the birth was precipitous. No other complications were recorded. The birth weight was 4 pounds 10 ounces. Good respiration was established within half a minute at which time the baby cried vigorously. He had good color and was at no time cyanotic or jaundiced. The following abnormalities were noted at birth: he had bilateral club feet (talipes equino varus on left and metatarsus varus on right) and a fine white line in a sagittal direction on the upper lip. Otherwise the physical examination was reported to be within normal limits. The mother was told that the placenta had been too small to nourish the fetus adequately but no pathological report is available. The baby was placed in an incubator for 2 weeks and remained in the newborn nursery at the hospital for 2 months. He sat at 9 months and walked only at 2 years of age, earlier development having been handicapped by casts and corrective shoes. He has never been heard to use any words. He has had the usual complement of inoculations and there have been no severe diseases. At age 2 he fell on his head and required two stitches; the incident had no sequelae. The first pediatric examination recorded in the hospital chart (age 2 years) revealed anomalies of the eyes in addition to the anomalies of feet and lip. There was a right ptosis and a convergent strabismus.

Family History

The family history is noncontributory. The subject is the second of three children: both siblings are intelligent and well. The mother is

[1] This paper was written while the author was a National Institute of Mental Health Career Investigator (Grant M-2921).

now divorced and lives with the subject's grandmother. A home visit revealed a warm and socially adequate climate.

Physical and Laboratory Examinations

When first seen by the author the child had been brought to a neurological service with a chief complaint of failure to develop speech. He was then 3 years and 9 months, of markedly small stature but with a head circumference normal for his chronological age (49.6 centimeters). Anomalies of the eyes had been corrected surgically. He had single palmar creases in both hands but no other mongoloid stigmata. The only other abnormal finding on examination was an enlarged heart; no murmurs were heard. His oral cavity was of normal configuration and he had no difficulties in chewing, swallowing, sucking, blowing, or licking. Laryngoscopy was negative. On radiological evidence his bone age was 2:8 according to Todd's standards. A skull series was normal. All laboratory tests were negative, and there were no signs of hypothyroidism or inborn errors of metabolism. An electroencephalogram was read as nondiagnostic though it was noted that activity in the right temporal area was less rhythmic than in the left.

Psychological Tests

Tests were performed at 4, 5, and 8 years. Merrill-Palmer, WISC, Bender Visual Motor Gestalt, and Leiter International Performance tests were used, the examinations being administered by three different psychologists and at two different clinics. The subject always related easily to the examiner and gave no signs of emotional disturbance or psychiatric disease. IQs were consistently in the 72–83 range but might have been consistently biased by the subject's inability to express himself verbally. He always gives an alert impression, reacts quickly to verbal instructions, and has always shown an adequate concentration span with little signs of distractibility or hyperactivity. At his most recent test no evidence of "organic" deficits was obtained although some "immaturity in his drawings" was noted. He is slightly retarded in his mental development but the deficit is definitely in the educable range and cannot explain his inability to speak.

Voice and Speech

The child's crying and laughter have always sounded normal. He is also able to make other noises, for instance short cough-like grunts accompanying his pantomimed communications. While playing alone he will readily make noises that sound somewhat like Swiss yodeling (though he has never had any experience with these sounds!) and which do not resemble any kind of vocalization heard among normal American children. (Samples of these sounds are reproduced in a documentary film.) When the author first saw him he appeared to have some difficulty in bringing his voicing mechanism under voluntary control. For instance, he was unable to make the pointer of the VU meter in an Ampex tape recorder jump by emitting grunts into a microphone even though he was fully aware of the logical connection between sound and deflection of the pointer and was fascinated by it. He would hold on to the microphone and move his head and lips toward it as if to prompt himself for the action; after a few futile attempts and with signs of rising frustration he would in desperation end up gesturing to the examiner's mouth inviting him to make the needle jump, or else simply resort to clapping his hands and accomplish his end this way. In recent years he seems to have learned to control his vocal apparatus to a greater extent. He has had speech training for a considerable length of time and can now repeat a few words after the speech therapist or his mother but the words are still barely intelligible and are never produced without the support from the speech correctionist or the mother (samples are reproduced in the film).

Some of the spontaneous sounds emitted by the patient at 4 years were analyzed spectrographically.[2] The spectrograms are grossly abnormal for a child of this age and resemble those of a neonate in a number of respects, such as the unsteadiness in the formant pattern, the intermittent bursts of nonharmonic overtones, and the almost random change in resonance distribution over the spectrum (Lenneberg, 1961). The spectrograms may be interpreted either as grossly immature or as evidence of a fixed central nervous system abnormality implicating the basic mechanisms for speech synergism.

From the patient's first visit to the clinic it

[2] A 5-page table giving a transcript of commands, questions, and statements with classification of subject's response and a figure showing frequency cycles per second have been deposited with the American Documentation Institute. Order Document No. 7317 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress; Washington 25, D. C., remitting in advance $2.50 for microfilm or $6.25 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

has been obvious that he had a normal and adequate understanding of spoken language. He has been seen more than 20 times since then and the finding of full comprehension has been confirmed by neurologists, psychologists, speech therapists, medical residents, and other hospital personnel. A number of tape recordings have been made of interviews including a visit to the patient's home. Most of the examinations were done without the presence of his mother. At one time a short series of instructions were tape recorded and transmitted to the patient through earphones. He followed the instructions without being able to see the examiner.

Demonstration Film

At age 8 years his capacity to comprehend was fully documented in a 16-millimeter sound film which is publicly available.[3] The film was not rehearsed and the examiner had been known to the patient by sight only. The demonstration includes the following items: ability to chew, swallow, and suck; sounds emitted while playing at age 4; tape recordings of mother's "conversation" with subject, recorded during a home visit; following commands and answering questions by nodding; a short story is told followed by questions on it which are couched in complex grammatical constructions such as the passive voice.

Interpretation

It is tempting to explain the patient's responses to verbal instructions by extralinguistic means. Perhaps he is merely responding to visual cues given by the examiner and has, in fact, not learned to understand English! Could children with his type of abnormality develop perceptual skills such as were observed in von Osten's famous horse, der kluge Hans, who supposedly could stamp his hoof in response to questions posed to him in German, but who, upon close examination by the psychologist Pfungst, had merely learned to observe the questioner, picking up minute motor cues related to posture and respiratory patterns which signaled to him whether to stop or to continue to stamp his hoof? There is direct evidence against this hypothesis. The child described can react to tape recorded instructions in the absence of any observer. Further, his responses do not merely consist of nodding but also of doing things which could under no circumstance be conveyed by inadvertent motor cues. In the film which documents the case, it is clear that the child frequently follows commands without looking at the examiner. Out of the 45 responses only 3 times was there vacillation between correct and incorrect answers and the last answer in each case is correct. On the other hand, there was no hesitation in the 3 instances when incorrect answers were given. There is no reason to assume that this child has superhuman ability to respond to visual cues instead of assuming that he has learned what every other child of his age has learned, namely, to understand English. Table 1 summarizes the child's performance. It is the result of a panel of three judges who scrutinized the film, viewing each command and its execution individually and with as many repetitions as was necessary in order to determine by unanimous agreement whether there was any likelihood of extralinguistic cuing.

Finally, we must consider the possibility that this patient had no understanding of syntactic connections but merely responded to key words in the commands and questions. This possibility is extremely unlikely in the face of his understanding of such sentences as "Take the block and put it on the bottle." "Is it time to eat breakfast now?" Was the black cat fed by the nice lady?"

[3] *The Acquisition of Language in a Speechless Child,* 16-millimeter sound film. Running time 18 minutes, distributed by Psychological Cinema Register, Pennsylvania State University.

TABLE 1

NUMBER OF SUBJECT'S CORRECT, INDECISIVE, AND INCORRECT RESPONSES

Type of response required	Classification of subject's response						
	Correct		Indecisive		Incorrect		
	No cue possible	Cue possible or certain	No cue possible	Cue possible or certain	No cue possible	Cue possible or certain	Total
Action	19	5	0	0	2	2	28
Yes-No nodding	2	11	0	3	0	1	17
Total	21	16	0	3	2	3	45

Diagnosis

The hospital's clinical diagnosis is multiple congenital anomalies, which, however, is a wastebasket classification and of no heuristic value. Certainly it does not explain the absence of motor speech. The two most common causes for this deficit, peripheral deafness and severe emotional disturbance, may be readily ruled out on clinical evidence. Nor may we assume that the patient's mental retardation is a sufficient cause since the degree of deficiency revealed in his psychological tests is not ordinarily accompanied by any marked speech deficit. Patients with an IQ as low as 25 to 35 have a wide repertoire of sounds and frequently use a vocabulary of 50 or more words. Some authorities would classify this patient as having congenital or developmental motor aphasia. I am not in favor of such a classification on terminological grounds. Aphasia has come to designate loss of speech in persons who had been fluent before the onset of disease or trauma. The condition occurs in children (Guttmann, 1942) as well as in adults and presents a symptomatology that is distinct from any developmental condition. However, the most important reason for rejecting the term aphasia for cases such as described here is that aphasia has traditionally been applied to cortical and subcortical lesions. The present case, on the other hand, gives every indication of an abnormality on a lower level, probably mesencephalic, because of the association with the ocular abnormalities and the discoordination as seen in the spectrograms. The term congenital anarthria better characterizes the condition. Psychological tests make cortical or subcortical damage also unlikely and his excellent understanding of language supports this view.

DISCUSSION

Failure to Learn to Understand Despite Babbling and Imitative Facility

The case reported makes it clear that hearing oneself babble is not a necessary factor in the acquisition of understanding; apparently, hearing oneself babble is also not a sufficient factor. In a language acquisition study on home-raised mongoloid children (research in progress) I have gathered empirical evidence that these children are excellent imitators who babble abundantly and freely. In all cases they have a speaking command of at least single word utterances; yet their understanding of complex commands and questions (such as used in the demonstration film) is frequently defective (usually if their IQ is about 50 or below). Here the vocal play is present and motivation is provided through interaction with parents, but it does not enable these children to overcome their inborn cerebral deficit.

Relationship between Speaking and Understanding

We must now pose the problem: if the secondary reinforcement provided by hearing oneself babble is neither necessary nor sufficient for an acquisition of understanding of language, could learning to speak be an entirely independent task, justifying a theory that applies to it alone, but not to the learning of understanding? We will see that the answer to this question is a qualified *no*.

The case presented here—by no means unique—is particularly dramatic because of the vast discrepancy between understanding and speaking; a similar phenomenon in more attenuated form is extremely common. Understanding normally precedes speaking by several weeks or even months. The discrepancy is regularly increased in literally all types of developmental speech disorders and is best illustrated in a condition known in the profession as Delayed Speech. Pertinent are also those children who have structural deformities in the oral cavity or pharynx and who produce unintelligible speech for years—sometimes throughout life—without the slightest impairment of understanding. Congenitally deaf children also learn to comprehend language in the absence of vocal skills. Understanding in all of these circumstances is definitely prior to and in that sense independent from speaking.

However, there is no clear evidence that speaking is ever present in the absence of understanding. Speaking is to be understood here as the production of utterances that are bona fide examples of a natural language (such as English) with presumptive evidence of autonomous composition of grammatically acceptable sentences. An empirical test of speaking without understanding might be as follows: a child acquires nothing but words that have no meaning to him (by blind imitation) and learns the formal principles governing the generation of sentences. He will now utter sentences out of context and irrelevant to situations by established common sense standards; he would also have to be demonstrably incapable ever to respond appropriately to commands formulated in natural language. My assertion is that such a condition has never been described as a congenital, developmental problem (Mark, 1962). (Adventiti-

ous conditions such as sensory aphasia in the adult are problems in *partial loss*, not *partial acquisition* of language and are therefore not relevant to this discussion. Nor could case reports of psychotic children be adduced reliably because even if their utterances are primarily echolalic, there are usually indications that the child does understand, at least at times, what is said to him. In fact, this is usually the basis of the therapy given them.) It is thus likely that the vocal production of language is dependent upon the understanding of language but not vice versa. Though there is no conclusive proof of this hypothesis, there are theoretical considerations that make this likely. In order to make this latter point clear, a few general remarks on the nature of grammar are indispensable. My discussion will lean heavily on Chomsky's (1957) and Chomsky and Miller's (1962) work.

Wherever the word *imitation* is used in connection with language learning it is assumed that subjects learn more than passive mirroring of sentences heard. The novel creation of sentences is a universally accepted fact. At first it was thought that this phenomenom could be accounted for by postulating that grammar simply reflected transitional probabilities between words. In this model the learning of grammar was thought to be like the learning of probabilities. Such learning has been demonstrated to occur for a great many mammals. If this model were acceptable, a child's exposure to certain contingencies should enable him to learn grammar in the absence of "understanding" the relationship between words. Chomsky has offered formal proof against left-to-right probabilistic models, and has shown how a different model can overcome some of the basic difficulties encountered by Markovian grammars. We shall not concern ourselves here with the mathematical detail but shall merely demonstrate by a few examples (most of them suggested by Chomsky) that grammar simply cannot be explained in terms of learned sequential contingencies, and that therefore the understanding and producing of sentences cannot be equated with probability learning.

Consider the following two strings of words, (*a*) colorless green ideas sleep furiously, (*b*) furiously sleep ideas green colorless. In terms of transitional probabilities they are indistinguishable. Both sentences can only occur in a zero order of approximation to English. However, we can discriminate between them from the point of view of grammaticality. Sentence *a* strikes us as a possible sentence, whereas *b* does not. The difference between *a* and *b* could not possibly be due to association by contiguity for obvious reasons (Miller, Galanter, & Pribram, 1960). Nor is the sequence of form class markers *-less, -s, -ly* the hallmark of grammaticalness as shown by the sentence, (*c*) friendly young dogs seem harmless, which is grammatical though it reverses the order of the markers.

In order to account for the difference in our perception of *a* and *b* we might try to see whether sequential contingencies do exist, but instead of on the level of words, on the level of parts of speech (actually proposed by Jenkins & Palermo, 1961). If this were so the essence of the transitional probability model would be saved and the principal underlying the formation of sentences would still be simple enough to allow of the possibility of speaking without understanding. Unfortunately this model is no more successful in accounting for "sentencehood" than any other Markovian device. If we compare the strings, (*d*) occasionally call warfare useless, (*e*) useless warfare call occasionally, we find that now *d* is perceived as more grammatical than *e* though the order of parts of speech in *d* is that of *b* above which we rejected as less grammatical. This example shows that the traditional categorization in terms of parts of speech is certainly not successful in reinstating a Markovian model.[4]

Mowrer (1960) has proposed that simple contiguity of words is sufficient explanation for the complex meaning that is conveyed by a sentence. Osgood (1957) on the other hand believes that grammatical order can be set aside by motivational factors and that with an increase of motivation word order would correlate with order of importance of words. Doubt is cast on both views, however, by comparison of sentences such as, (*f*) the fox chases the dog, (*g*) the dog chases the fox, which clearly make either position untenable. A child whose task it is to learn to produce sentences such as *f* and *g* in the appropriate physical environment must necessarily learn some principal of concatenation which goes *beyond* recognition of contiguity. It would be tempting to maintain that the principle to be learned is something like "First noun phrase is the actor or subject in a sentence." Yet, the patterning involved must be more complex and in a sense more abstract still, for

[4] Nor is there any hope that eventually more efficient categories might be discovered since Chomsky's (1957) criticism is leveled against all finite state sentence generating devices. Compare his argument concerning mirror-image languages, and Miller's (1960) argument.

even a preschool child would understand the sentence, (*h*) the fox is chased by the dog, as belonging to *g* and not to *f*. To explain this phenomenon we would now have to postulate that in addition to the principle above, the child would also have to learn that the presence of the morphemes *is, -ed, by* signals a reversal of the original principle. But sentence (*i*) the fox is interested by virtue of his nature in chasing the dog, eliminates this possibility because this sentence is understood as similar in meaning to *f* instead of *g* despite the presence of morphemes *is, -ed, by* occurring in essentially the same sequential order as in *h*. In other words, if we must learn to compose sentences that conform to English grammar, we can only learn to apply the structural principles of sentence formation by first learning to understand and to group sentences in accordance with similarity or differences in meaning. It is not possible to explain the grammatical phenomena demonstrated in the test sentences above by assuming that the entire sentence is in one way or another associated with the complex natural situation to which it refers. Sentences without any referential meaning such as, (*j*) A v's C, (*k*) C v's A, (*l*) C is v'ed by A, (*m*) A is v'ed by C, can easily be grouped in terms of similarity and dissimilarity of meaning. Therefore, the word *meaning* in this context refers to grammatical understanding and not to an association between a symbol and a physical stimulus.

Obviously, we do not yet have a satisfactory model that might explain how grammar is learned. All we can say is that a child learns to produce novel sentences after hearing a number of utterances which were produced by formal laws of generation in addition to a number of other determining factors: he must be able to abstract the formal laws through observation (or be equipped to accept sentences as an input and recognize invariant patterns of complex relationship) before he can apply them to the production of new sentences. It is particularly important to realize that what the child learns during acquisition of grammar is the peculiar formal relationship that obtains between a number of different grammatical patterns, i.e., the relationship between active-passive, declarative-interrogative-negative, and many other similar relationships called *rules of transformation* (Chomsky, 1957; Harris, 1957). It is this latter ability—and only this—that enables any speaker of English to group the Sentences *f* through *i* according to similarity in meaning. The psychological process involved here is more similar to the operations by which we know that:

$$w\ (\sqrt{w} \times \sqrt{w})$$

and

$$\frac{w^6}{w^4}$$

can both be represented by w^2 than to an operation by which we know that the symbol S_4 is next in line to be generated after a train of symbols S_1, S_2, S_3, has been produced. This point is illustrated by Chomsky's examples of structural ambiguity. There are rules of transformations which convert sentences of the form (*n*) one visits relatives, (*o*) relatives are visiting, into a phrase, (*p*) visiting relatives. Because of the two different transformational origins this phrase is ambiguous and, when used in a sentence, may render that entire sentence ambiguous: (*q*) visiting relatives can be a nuisance.

The conclusion of this discussion is that *knowing a language* may be, and ordinarily is, manifested by two distinct behavioral manifestations: understanding and speaking. Upon careful analysis, however, both of these manifestations depend upon the application and use of a single set of grammatical rules; in the case of understanding, the rules are applied to the analysis, i.e., processing and organizing of input data; in the case of speaking, the same rules are applied to the organization of output data or responses. In the process of language learning, the acquisition of grammatical rules must occur first in connection with analysing incoming sentences; then with producing outgoing sentences. The most important point here is, however, that *knowing* a natural language is dependent upon the *acquisition of a single set of organizing principles* and that this set of principles is merely reflected in understanding and speaking but is not identical with these skills.

SUMMARY

A case was presented, typical of a larger category of patients, where an organic defect prevented the acquisition of the motor skill necessary for *speaking a language*, but evidence was presented for the acquisition of grammatical skills as required for a complete *understanding of language*. Theories on the acquisition of speech and language must account for both motor and grammatical skills. Present theories assert that babbling, hearing oneself vocalize, and imitation are the cornerstones of speech development. These phenomena primarily relate to the development of the motor skills involved which, however, never develop in isolation, i.e.,

without simultaneous acquisition of grammatical skills. The case presented together with the language deficit in certain Mongoloids clearly shows that babbling, hearing oneself talk, and imitation are neither sufficient nor necessary factors in the acquisition of grammar and since the motor skills alone are never shaped into "speaking without grammar," i.e., parroting without understanding, it is concluded that the present theories are inadequate.

REFERENCES

CHOMSKY, N. *Syntactic structures*. Hague: Mouton, 1957.

CHOMSKY, N., & MILLER, G. A. Introduction to the formal analysis of natural languages. In D. Luce, E. Galanter, & R. Bush (Eds.), *Mathematical psychology*. 1962, in press.

GUTTMANN, E. Aphasia in children. *Brain*, 1942, **65**, 205–219.

HARRIS, Z. S. Co-occurrence and transformation in linguistic structure. *Language*, 1957, **33**, 283–340.

JENKINS, J. J., & PALERMO, D. S. Mediation processes and the acquisition of linguistic structure. Paper read at SSRC conference in Cambridge, Massachusetts, October 1961.

LENNEBERG, E. H. Speech as a motor skill with special reference to nonaphasic disorders. Paper read at SSRC conference in Cambridge, Massachusetts, October 1961.

MARK, H. J. Elementary thinking and the classification of behavior. *Science*, 1962, **135**, 75–87.

MILLER, G. A. Plans for speaking. In G. A. Miller, E. Galanter, & K. H. Pribram (Eds.), *Plans and the structure of behavior*. New York: Holt, 1960. Pp. 139–158.

MILLER, G. A., GALANTER, E., & PRIBRAM, K. H. *Plans and the structure of behavior*. New York: Holt, 1960.

MOWRER, O. H. *Learning theory and the symbolic processes*. New York: Wiley, 1960.

OSGOOD, C. E. Motivational dynamics of language behavior. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1957*. Lincoln: Univer. Nebraska Press, 1957. Pp. 348–424.

(Early publication received April 5, 1962)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 426–427

CRITIQUE AND NOTES

TWO ANXIETY FACTORS IN PSYCHIATRIC PATIENTS

ARNOLD H. BUSS

University of Pittsburgh

Eysenck (1961) states that there are two factors in anxiety, "one factor of autonomic overreactivity, and one characterized by the conditioned anxiety responses" (p. 21). This hypothesis received confirmation from a study by Hamilton (1959), who interviewed anxiety neurotics and rated various anxiety symptoms. The intercorrelations of the ratings yielded a bipolar factor consisting of muscular tension, fears, insomnia, apprehension, difficulty in concentrating, and depression versus autonomic symptoms, gastrointestinal complaints, genitourinary complaints, respiratory complaints, and cardiovascular complaints. Tension and worry may be regarded as conditioned anxiety, and the various physical complaints may be regarded as symptomatic of autonomic reactivity.

Hamilton's (1959) subjects were patients diagnosed as anxiety-neurotics. It would be of considerable interest to discover whether similar anxiety factors (tension, worry versus autonomic symptoms) would be found in a sample of psychiatric patients of varying diagnoses. If these two factors emerged from ratings by others of the patient's symptoms, it would lend greater generality to Eysenck's (1961) hypothesis.

METHOD

The subjects were 73 patients, 30 men and 43 women, in a psychiatric hospital. Diagnoses varied greatly, and those with organic brain damage were excluded. Each subject was interviewed by the writer with three observers[1] in the room. At first the interview was open-ended and the questions general ("What has been bothering you?"). Later the questions were more specific in order to insure that the same areas were explored in each subject. The interviews lasted approximately 15 minutes, after which ratings were made independently by the interviewer and the three observers.

The following categories were rated on a five-point scale:

1. Physiological concomitants (observed)—Sweating, flushing, shallow breathing, excessive swallowing
2. Physiological concomitants (reported)—Same as Number 1, plus heart palpitations
3. Subjective feelings—Reported discomfort, "butterflies in the stomach"
4. Somatic complaints—Reported aches and pains (frequency, chronicity, and intensity)
5. Distractibility—Attention lapses and poor performance on subtracting sevens from 100 (which was done in the final phase of the interview)
6. Restlessness—Squirming in the chair, aimless movements, tremors, and tics
7. Worry
8. Muscular tension—Reported tightness, cramps, and "being tied up in knots"

[1] The observers were Marc Baer, Ann Durkee, and Morton Wiener.

RESULTS

The pooled ratings for the categories were intercorrelated, and the matrix is presented in Table 1. The correlations range from near 0 to .63, with most being in the thirties and forties. Evidently, various aspects of anxiety, while positively related to each other, are not highly correlated with each other; anxiety does not appear to be univocal.

A factor analysis of the correlations (Thurstone's centroid method) yielded two factors, which were rotated to simple structure. The factor loadings are presented in Table 2. The categories loading on Factor I are Physiological concomitants (both observed and reported), Subjective feelings, Somatic complaints, and Distractibility. Except for Distractibility, this factor clearly refers to the physiological symptoms of anxiety that Eysenck (1961) has labeled autonomic overreactivity. Factor II is defined by Restlessness, Worry, and Muscular tension; these motor and ideational aspects of anxiety clearly

TABLE 1

CORRELATION MATRIX: COMPONENTS OF ANXIETY

Component	2	3	4	5	6	7	8
1. Physiological concomitants (observed)	.30	.41	.24	.36	.43	.14	.19
2. Physiological concomitants (reported)		.63	.45	.38	.36	.32	.53
3. Subjective feelings			.42	.29	.28	.46	.32
4. Somatic complaints				.32	.28	.12	.42
5. Distractibility					.36	.03	.43
6. Restlessness						.46	.61
7. Worry							.53
8. Muscular tension							

TABLE 2

FACTOR ANALYSIS: COMPONENTS OF ANXIETY

Component	I	II
1. Physiological concomitants (observed)	.43	.00
2. Physiological concomitants (reported)	.55	.08
3. Subjective feelings	.62	−.02
4. Somatic complaints	.46	.00
5. Distractibility	.45	.00
6. Restlessness	.01	.62
7. Worry	−.05	.45
8. Muscular tension	.15	.46

fall under the heading of what Eysenck has called conditioned anxiety.

Thus the results of this study are consistent with those of Hamilton (1959) in finding two anxiety factors: autonomic overreactivity and conditioned anxiety. The single discrepancy was that Distractibility loaded on the autonomic factor; if this finding is confirmed by further research, it might be speculated that physiological symptoms interfere with attentiveness. The present study extends Eysenck's (1961) hypothesis to psychiatric inpatients, and it may be concluded that his two-factor approach to anxiety, which has led to fruitful laboratory research, is also clinically meaningful.

REFERENCES

EYSENCK, H. J. *The handbook of abnormal psychology*. New York: Basic Books, 1961.

HAMILTON, M. The assessment of anxiety states by rating. *Brit. J. med. Psychol.*, 1959, **32**, 50–59.

(Received July 18, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 427–428

FURTHER VALIDATION OF THE ALPERT-HABER ACHIEVEMENT ANXIETY TEST

WILLIAM N. DEMBER, FRANK NAIRNE, AND FRANCIS JOSEPH MILLER [1]

University of Cincinnati

This note reports some further validation data on the recently devised test anxiety scale described in this journal by Alpert and Haber (1960). The Alpert-Haber scale was designed to measure anxiety experienced in conjunction with test taking; it differs from the Mandler-Sarason (1952) scale in that it distinguishes between anxiety that is reported by the respondents as debilitating and anxiety reported as facilitating test performance. Each type of anxiety is measured in a separate subset of items of a 19-item scale.

To summarize briefly a considerable amount of data, Alpert and Haber (1960) report that the two subscales are negatively correlated with each other ($r = -.48$) and that each of the scales correlates (r's in the .20s) with: performance on the SAT; score on a midterm test, final exam grade, and course grade in introductory psychology; and (r's in the .30s) with overall grade-point average. Multiple correlations, predicting both grade-point average and final exam grade from a combination of SAT and either of the anxiety scores, ranged from .29 to .58.

In two studies, we have collected data on measures comparable to those above. The two populations from which our subjects were drawn differ from each other and from the population (Stanford University students) sampled by Alpert and Haber (1960); and our performance measures are somewhat different from theirs. Thus, our studies not only provide an opportunity for confirmation of the data reported by Alpert and Haber, they also bear on the generality of the earlier results. Finally, we have some evidence on the relative usefulness of the scale for male and female subjects.

STUDY I

The subjects in the first study were 39 male students enrolled in an experimental psychology class in a Catholic seminary in Norwood, Ohio. Their age range was from 20 to 30 years, with a median of 22. For each subject the following performance data were obtained: score on the School and College Ability Test (SCAT); score on the midterm test (MTT) in experimental psychology; overall grade-point average (GPA) earned in the previous semester. The Alpert-Haber question-

[1] The authors are grateful to Jane Blumberg and Robert McCullers for their participation in this research.

TABLE 1

INTERCORRELATIONS BETWEEN ANXIETY AND PERFORMANCE MEASURES IN STUDY I

Measure	AAT−	SCAT	Midterm test	Grade-point average
AAT+	−.65**	+.27*	+.43**	+.52**
AAT−		−.35*	−.46**	−.37*
SCAT			+.39**	+.61**

* $p < .05$, $df = 37$, one-tailed test.
** $p < .01$, $df = 37$, one-tailed test.

naire was administered by the instructor (FJM) as part of the regular course work, and the subjects were assured that the information would be used for research purposes only. For each subject a facilitating (AAT+) score and a debilitating (AAT−) score was determined.

Pearson product-moment correlation coefficients were computed between the two anxiety measures, and between each of them and the three performance measures. In addition, correlation coefficients were computed between SCAT and MTT and between SCAT and GPA; these were needed for the multiple correlations described below. The results are given in Table 1. All of the correlation coefficients are significantly different from zero, and in general are considerably higher than those reported by Alpert and Haber (1960) for comparable measures.

Multiple correlations were also computed between each of the anxiety measures and the SCAT as a pair of predictor variables and each of the academic performance measures as criteria. These *R*s are given in Table 2.

STUDY II

Alpert and Haber (1960) give results only for male subjects. Our second study was carried out mainly for the purpose of comparing the usefulness of the scales for male and female subjects. There is reason to expect that such a comparison might reveal important differences (see, for example, Carlson & Carlson, 1960).

The subjects in this study were 25 male and 39 female students enrolled in an introductory psychology class in the University of Cincinnati. The anxiety questionnaire was administered by the instructor,[2] at the beginning of a regular class session. The only measure of academic performance examined for these subjects was their final grade in the introductory psychology class.

TABLE 2

MULTIPLE CORRELATIONS PREDICTING MIDTERM TEST (MTT) AND GRADE-POINT AVERAGE (GPA) FROM A COMBINATION OF SCHOLASTIC APTITUDE (SCAT) AND ANXIETY

Measure	Performance measure	
	MTT	GPA
SCAT and AAT+	.51	.71
SCAT and AAT−	.52	.63

TABLE 3

INTERCORRELATIONS BETWEEN ANXIETY AND PERFORMANCE MEASURES IN STUDY II

Measure	AAT−	Course grade
Males ($N = 25$)		
AAT+	−.65**	+.59**
AAT−		−.34*
Females ($N = 39$)		
AAT+	−.42**	+.06
AAT−		−.23

* $p < .05$, one-tailed test.
** $p < .01$, one-tailed test.

Correlation coefficients were computed, separately for male and female subjects, between the two anxiety measures and between each of these and course grade. The results are given in Table 3.

For the female subjects only the correlation between AAT+ and AAT− is significantly different from zero. The correlation of −.23 between AAT− and course grade for the females is lower ($p < .05$) than the correlation of −.34 for the males. The correlations for the male subjects are all significantly different from zero and generally higher than those obtained by Alpert and Haber (1960).

CONCLUSION

The two studies together lend confirmation and generality to the original Alpert-Haber (1960) data. Results of the second study also suggest a limitation on the applicability of the scale: that it is not as useful for female as for male subjects. The high correlations obtained with our male subjects, as compared with those obtained with the Stanford students used by Alpert and Haber, suggest that variables other than sex might also influence the usefulness of the scale. It would be premature to speculate on the identity of those variables.

[2] A. B. Kristofferson administered the anxiety questionnaire and provided the information on course grades.

REFERENCES

ALPERT, R., & HABER, R. N. Anxiety in academic achievement situations. *J. abnorm. soc. Psychol.*, 1960, **61**, 207–215.

CARLSON, E. R., & CARLSON, RAE. Male and female subjects in personality research. *J. abnorm. soc. Psychol.*, 1960, **61**, 482–483.

MANDLER, G., & SARASON, S. B. A study of anxiety and learning. *J. abnorm. soc. Psychol.*, 1952, **47**, 166–173.

(Received July 18, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 429–431

THE EFFECTIVENESS OF SOCIAL REINFORCEMENT FOLLOWING TWO CONDITIONS OF SOCIAL DEPRIVATION[1]

HAROLD W. STEVENSON AND RICHARD D. ODOM

University of Minnesota

Gewirtz and Baer (1958a, 1958b) have demonstrated that the effectiveness of social reinforcement in modifying children's performance is increased by a brief period of social isolation preceding the experimental task. The results are interpreted as indicating that the effectiveness of a social reinforcer is increased by its own deprivation.

Another interpretation of these findings is possible. Isolation resulted in the deprivation of not only social but also of other types of stimuli. The children, therefore, may have been subjected to general stimulus deprivation rather than merely to deprivation of social stimuli. As a consequence, the effectiveness of social stimuli as reinforcers may be reduced when other types of stimuli are available during the deprivation period. One approach to determining the degree to which the increased effectiveness of social reinforcement is due specifically to the deprivation of social stimuli would be to isolate subjects in a setting providing a wide variety of other types of stimuli and to compare their performance with that occurring in a condition where other subjects were isolated without such stimuli.

This study investigates the effect of social reinforcement on children's performance in a simple operant task following three conditions: a control condition with no isolation; a "toy" condition in which the subject is left alone in a room filled with unusual and interesting toys with which he can play freely for a 15-minute period; an isolation condition in which the subject is left alone in the room with no toys for 15 minutes.

METHOD

Subjects

The subjects were 30 boys and 30 girls attending kindergarten, first, and second grades of the Laboratory School of the University of Minnesota.[2] The subjects were above average in intelligence and socioeconomic status.

Apparatus

Experimental task. The apparatus was a red rectangular box, 22 inches long, 16 inches wide, and 12 inches deep. Two 8-inch square sunken bins were located on the top of the box. The right bin was covered by a Masonite plate with six 5/8-inch holes randomly placed on its surface. The floor of the right bin was covered with foam rubber and sloped toward a small aperture. As a marble passed through the aperture it activated a microswitch which in turn activated an electric counter. The left bin was filled with approximately 600 marbles of different colors.

The apparatus was placed on a child-sized table in an 8 foot × 7 foot room. The only other furnishing in the room was a chair for the experimenter behind the apparatus.

Pretraining task. The pretraining room was 10.5 feet long and 10 feet wide and was located two rooms away from the experimental room. It contained two tables, two chairs, a sink, windows with drapes drawn, and a one-way vision mirror.

For subjects in the Toy condition the tables and chairs were filled with toys. The toys included a drumming clown, a drinking rabbit, a rocking-telephoning bear, a balloon-blowing bear, a typewriting doll, a bouncing bird, a butterfly-chasing duck, a spiral shooting gun, a large stuffed bear, a stuffed rabbit, a beach ball, and a xylophone. All except the last seven toys were powered by batteries and were equipped with 2-foot extension cords with switches which could be easily operated by the children.

Procedure

The subjects were randomly divided into three groups with an equal number of boys and girls in each group. For the two pretraining groups a 15-minute period in the pretraining room preceded the experimental task. In the Toy group the pretraining room contained the toys described above. In the Isolation group no toys were present. The preliminary 15-minute period was eliminated for the control group.

The subjects were obtained individually by the experimenter from their classroom. While the experimenter escorted the subject to the laboratory room, the experimenter attempted to gain rapport with the subject by talking about school and related

[1] This study was supported by a grant from the National Institute of Mental Health (Grant M-3519) to the first author.

[2] The writers wish to thank the staff of the University Laboratory School for their cooperation in this study, and Leone MacLean, who served as the experimenter.

matters. The experimenter was a female elementary teacher trained for the experiment.

Pretraining task. For subjects in the Toy and Isolation groups the experimenter approached the experimental room and told the subject,

> Now the game is in that room right there, but I have to get it ready. You wait in here and I will come and get you when the game is ready.

The experimenter walked with the subject into the pretraining room.

For the Isolation group the experimenter told the subject,

> As soon as I get the game ready I will come back and get you. You wait in here until I get the game ready.

The experimenter went out of the room leaving the door partly open.

For the Toy group the experimenter told the subject,

> I have a lot of toys in here and you can play with them while you are waiting. Play with any of the toys you want to and have fun.

The experimenter demonstrated how one of the mechanical toys worked. Again, the experimenter left the door partly open.

As the experimenter left the pretraining room she started a stopwatch and went to a hallway from which she could observe the subject's behavior through the one-way mirror of the pretraining room. At the conclusion of the 15-minute period the experimenter returned to the pretraining room and told the subject,

> The game is ready. You can come with me now.

Experimental task. The experimenter took the subject to the experimental room and said,

> This is a game called Marble-in-the-Hole. I'll tell you how to play it. See these marbles. Well, they go in these holes like this. [The experimenter demonstrated.] You pick up the marbles one at a time and put them in these holes. I'll tell you when to stop. Remember, pick them up one at a time, and I'll tell you when to stop.

The task was concluded after 7 minutes.

After an average of every 15 responses the experimenter made a supportive comment about the subject's performance. The comments were delivered after every tenth, fifteenth, or twentieth response according to a prearranged random schedule. Six statements were used in a random order: Fine, Very good, That's fine, You're really good, Good, and Swell.

The experimenter counted the subject's responses in order to follow the reinforcement schedule, but the responses were also recorded electrically for each successive minute.

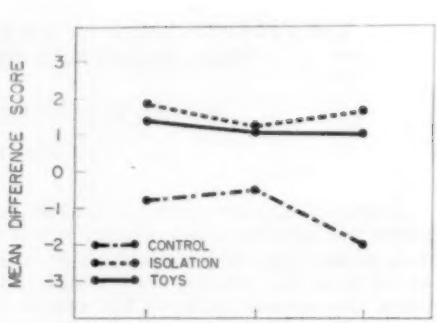

FIG. 1. The average difference score for each group for 2-minute intervals following the first minute.

RESULTS

The score used in the analysis of the results was a difference score obtained by subtracting the number of responses made during the first minute of the task from that for each successive minute. This procedure was adopted to reduce the effect of individual differences in response rate.

The average number of responses during the first minute of the game was 25.8. A 2×3 analysis of variance with entries for sex of subject and condition resulted in no significant main effects and in a nonsignificant interaction (Fs < 1.00). Performance during the first minute of the game may be assumed to be approximately the same in all conditions.

The major results of the study are summarized in Figure 1, which presents the average difference score obtained by each group for 2-minute intervals after the first minute of the task. The highest scores occur in the Isolation group and the lowest in the Control group. The scores of subjects in the Toy group were slightly below those of the Isolation subjects. A facilitating effect, which persisted throughout the task, was evident in the Isolation and Toy groups. The subjects in the Control group showed a decrement in performance after the first minute of the task which increased during the last 2 minutes.

These data were subjected to a 2×3 analysis of variance. The F associated with condition was significant at the .05 level ($F = 3.76$, $df = 2/54$). The F associated with sex of subject was not significant ($F < 1.00$), nor was the interaction term ($F = 2.64$, $df = 2/54$, $p > .05$).

DISCUSSION

The toys appeared to be very attractive to the subjects and they played with them with enthu-

siasm. When no toys were present the subjects sat quietly, sighed, peered around the room, and showed other indications of impatience and boredom. Although some subjects in the Toy group showed some of these types of behavior towards the end of the 15-minute period, the behavior of the two groups was in general quite dissimilar.

The greater effectiveness of social reinforcement in modifying performance in the Isolation group compared to the Control group corroborates the results of Gewirtz and Baer (1958a, 1958b). The greater effectiveness of a female experimenter with boys than with girls found by Gewirtz and Baer was not obtained; however, older subjects were used in the present study.

The presence of other stimuli during the isolation period did not result in a significant decrease in the effectiveness of social reinforcement over that obtained when no toys were present. Although the difference scores for the subjects in the Toy group are consistently below those in the Isolation group, the difference is not significant ($t < 1.00$). Perhaps with a larger number of subjects or with a more sensitive measure of performance a significant difference would emerge. For the present, however, it may be tentatively concluded that the greater effectiveness of social reinforcement following isolation is primarily dependent upon deprivation of social stimuli, rather than on more general stimulus deprivation.

SUMMARY

Three groups of elementary school children were presented with a simple operant task in which the insertion of marbles into a series of holes was verbally reinforced. The first group of subjects played alone for 15 minutes in a room containing a variety of interesting toys. The second group was left alone for 15 minutes with no toys. A control group received no pre-experimental treatment. Analysis of variance indicated a significant difference among conditions. The Isolation group showed the greatest increase in response rate after the first minute of the task and the Control group the lowest. The performance of the Toy group did not differ significantly from that of the Isolation group. The results are interpreted as tentatively supporting the hypothesis that the increased effectiveness of social reinforcement following isolation is primarily dependent upon the deprivation of social stimuli rather than upon more general stimulus deprivation.

REFERENCES

GEWIRTZ, J. L., & BAER, D. M. Deprivation and satiation of social reinforcers as drive conditions. *J. abnorm. soc. Psychol.*, 1958, **57**, 165–172. (a)

GEWIRTZ, J. L., & BAER, D. M. The effect of brief social deprivation on behaviors for a social reinforcer. *J. abnorm. soc. Psychol.*, 1958, **56**, 49–56. (b)

(Received July 20, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 431–438

A CROSS-VALIDATION OF TWELVE MMPI INDICES OF HOSTILITY AND CONTROL

EDWIN I. MEGARGEE AND GERALD A. MENDELSOHN[1]

University of California

Dahlstrom and Welsh's (1960) recently published MMPI handbook contains an appendix listing 213 scales, indices, and scoring procedures for the MMPI. This proliferation of scales has produced nearly one scale for every two MMPI items. Each of these scales is supposed to measure validly some meaningful facet of human personality ranging from "hypochondriasis" to "success in baseball." While the 10 original scales and some of the new ones were derived by careful empirical study, not all authors have been equally thorough in establishing the validity of their instruments. Consequently there is considerable variability in the confidence with which the 213 scales may be used.

The clinician or researcher who wishes to measure some aspect of personality by using one of these scales often has little information about their real meaning or usefulness. A search of the literature generally reveals little data on which to base a decision, for the amount of

[1] The authors wish to express their appreciation to Lorenzo Buckley, Chief Probation Officer of Alameda County, California, John Clausen of the Institute of Human Development, and their respective staffs for providing the data for this study. In addition, the writers are grateful to Jack Block for his help in analyzing the data of this study and the University of California Computer Center for donating time on the IBM 704.

cross-validation is too often nonexistant or inadequate. Frequently the only information available is that which Dahlstrom and Welsh (1960) have gleaned from unpublished sources, or the title which the scalemaker has given to his instrument. The question naturally arises as to how valid these scales are and whether their publication represents progress or merely additional noise in the field. The present study, in which an attempt was made to cross-validate the scales and indices relating to a single aspect of behavior, focuses on this question.

The particular aspect of behavior studied, assaultive acting out, was chosen because of the need of the senior author to measure aggressive potential in candidates for probation. Such measurement would be of great value in deciding whether a criminal should be incarcerated or could be safely returned to the community. Since a number of MMPI scales purport to measure such traits as hostility, impulsivity, and control, a study was undertaken to select those best suited to the above purpose. This involved a cross-validation of the relevant scales using groups

TABLE 1

DESCRIPTION OF SCALES AND INDICES ASSESSING HOSTILITY AND CONTROL

Symbol	Name	Author	Description
Ap	Adjustment to Prison	Panton, 1958	36 items, empirically derived from protocols of 56 nonadjusted male prisoners versus 72 adjusted inmates. Adjustment is defined as absence of severe infractions such as fighting, assaulting a guard, or refusing to obey orders. Scale thus derived was later validated and refined using two adjusted samples totaling 177 subjects and three nonadjusted samples totaling 204 subjects. Scales held up with $p \leq .01$ on these cross-validations.
Hc	Hostility Control	Schultz, 1954	34 items, empirically derived. One hundred nineteen Veterans Administration therapy patients were rated by their therapists on the adequacy of the methods they used to deal with hostility. The protocols of the top and bottom 25% were used in the item analysis. Obtained items later cross-validated on therapists' ratings in a college sample.
Ho	Hostility	Cook & Medley, 1954	50 items. The MMPI protocols of those teachers in Minnesota who were rated in the top and bottom 8% of the teachers in the state on the basis of the Minnesota Teacher Aptitude Inventory were compared. Fifty of the discriminating items were selected by five judges as measuring hostility. McGee (1954) obtained a significant correlation with hostility as reflected in Szondi sorts.
Hv	Overt Hostility	Schultz, 1954	14 items. Same derivation as Hc except therapists rated the patients' frequency of expression of overt hostility.
Hy – S	Inhibition of Aggression	Harris & Lingoes, 1955	7 items in the *Hy* scale which, on rational examination by the authors, appear to reflect this quality.
Im	Impulsivity	Gough, 1960	21 items. This consists of those items on the CPI "self-control" scale which are common to both tests. People scoring high on Im are ". . . impulsive, shrewd, excitable, irritable, self-centered, and uninhibited. Aggressive and assertive" (Gough, 1960, p. 12). Cross-validations yield positive correlations with ratings of impulsivity in normal samples.
Jh	Judged Manifest Hostility	Siegel, 1956	47 items. One hundred ten items deemed to indicate hostility were submitted to five judges who chose items reflecting manifest hostility as defined by Veterans Administration nomenclature. Items on which there was 80% agreement were used in the scale.
Eo	Ego Overcontrol	Block, 1955	23 items. Empirically derived from samples of 200 males tested at the Institute of Personality Assessment and Research, Berkeley, California, designated Ec-3 by Block, Eo by Dahlstrom and Welsh.
Nu	Neurotic Undercontrol	Block, 1955	33 items. Dahlstrom and Welsh (1960) quote Block as stating this measures ". . . susceptibility of the personality structure to stress and anxiety in which the anxiety is handled by impulsive acting out, aggression, and erratically directed hyperactivity . . ." (p. 302, source not cited). Block* reports this scale is effective only if the subject is both neurotic and undercontrolled.
Bc	Bimodal Control	Block, 1955	48 items. "By bimodal control is meant an inconsistency in controls and expressive modes of behavior . . . . In some behavioral areas he acts in an overcontrolled manner and in other areas in an undercontrolled manner" (Block, 1953, p. 1).
AHI	Active Hostility Index	Welsh & Sullivan, 1952	The sum of *T* scores on *Pd* and *Ma*. Means and *SD*s of various samples reported in Dahlstrom and Welsh (1960).
FTI	Frustration Tolerance Index	Beall & Panton, 1957	A ratio of the *T* scores on four clinical scales: $FTI = \frac{Pd + Ma}{D + Hy}$

* J. Block, personal communication, 1961.

TABLE 2

MEANS AND STANDARD DEVIATIONS OF ALL GROUPS ON ALL SCALES

Scale	Group I Extremely assaultive criminals		Group II Moderately assaultive criminals		Group III Nonassaultive criminals		Group IV Noncriminals	
	$\bar{X}$	*SD*	$\bar{X}$	*SD*	$\bar{X}$	*SD*	$\bar{X}$	*SD*
L	4.786	3.215	5.800	2.814	4.760	2.905	2.760	1.666
F	5.429	3.502	3.960	3.034	6.080	4.609	3.400	1.936
K	15.429	5.417	17.800	5.492	15.320	5.793	14.680	3.287
Hs	7.429	5.680	4.440	5.628	5.000	5.260	5.040	3.910
Hs +.5*K*	15.357	5.692	13.600	5.132	12.880	5.011	12.600	4.472
D	21.429	3.857	20.240	5.761	23.240	5.925	17.600	4.103
Hy	21.857	5.573	23.360	6.343	20.120	5.380	19.680	5.942
Pd	20.786	3.378	19.080	4.847	20.480	4.104	14.760	4.549
Pd +.4*K*	26.929	3.562	26.200	3.948	26.640	4.221	20.560	4.976
Mf	24.571	4.291	25.000	5.173	23.160	5.080	25.120	4.106
Pa	10.857	3.592	9.560	2.873	10.200	3.663	8.640	2.447
Pt	12.857	8.698	9.160	8.764	11.160	7.386	9.920	5.385
Pt + 1*K*	28.286	5.928	26.960	5.624	26.480	5.205	24.600	5.307
Sc	12.929	9.311	9.680	9.616	11.000	9.828	6.240	3.282
Sc + 1*K*	28.357	6.570	27.480	7.512	26.320	7.728	20.840	4.327
Ma	16.571	4.637	15.800	4.388	16.720	4.147	17.440	4.565
Ma +.2*K*	19.571	4.256	19.440	3.852	19.240	3.951	20.280	4.569
Si	28.714	8.081	25.440	8.021	28.840	10.535	24.480	10.481
AHI	125.714	16.406	123.800	15.777	122.920	17.445	112.520	21.028
FTI	1.051	.179	1.041	.170	1.036	.205	1.062	.207
Ap	12.786	2.607	14.640	2.215	13.560	3.513	15.320	2.625
Hc	7.714	4.214	6.800	3.674	8.280	4.306	5.600	2.273
Ho	20.714	8.973	15.440	8.451	19.480	9.592	16.480	5.775
Hv	4.429	1.785	3.800	1.915	4.040	1.946	3.960	1.695
Hy − 5	3.286	1.729	4.160	0.850	3.000	1.500	3.400	1.190
Im	6.643	4.448	5.560	3.465	6.680	3.350	7.200	1.936
Jh	14.786	6.997	10.800	6.351	13.360	7.012	13.120	4.428
Eo	13.000	4.132	13.840	3.091	13.360	2.643	10.520	3.293
Nu	12.571	6.186	10.360	4.923	12.520	5.051	12.120	3.100
Bc	22.357	9.565	21.240	10.829	23.080	8.514	20.320	6.656

known to differ, on the basis of their past behavior, in regard to these traits. Criminal assaultiveness was used as an operational definition of hostility and lack of control. The assumption was made that any useful scale of hostility or control should discriminate between criminals convicted of highly aggressive crimes and those criminals whose records showed no history of overt violence.

PROCEDURE

Scales

Ten scales and two scoring indices for the MMPI were selected. In each case either the name or the description of the scale indicated that a trait was measured which may be assumed to be related to the expression or control of hostility. A description of these scales appears in Table 1.

Subjects

The MMPI protocols of three groups of men referred to the Alameda County, California, Probation Department Guidance Clinic during the last 3 years were selected from the clinic files.

Group I, the extremely assaultive group, consisted of 14 men convicted of violation of one or more of the following sections of the California Penal Code (State of California, 1959): 187, Murder, second degree; 192.1, Voluntary manslaughter; 203, Mayhem; and 245, Assault with a deadly weapon. Their ages ranged from 20 to 51 with a mean of 30.

Group II, the moderately assaultive group, consisted of 25 men randomly selected from those referred to the clinic after having been convicted of Section 242, Battery, defined in the Penal Code (State of California, 1959) as ". . . the unlawful use of force or violence on the person of another. . . ." Their ages ranged from 18 to 68 with a mean of 36.

Group III was a control group consisting of 25 men randomly selected from all those who had been referred during the last 3 years after conviction of a nonassaultive crime. In order to ensure that the members of this group would be nonassaultive, anyone convicted at any time of one of the violent crimes mentioned above, as well as men who were

convicted of other lesser charges which have an assaultive component, were excluded from consideration. The crimes thus excluded were such offenses as attempted murder, assault, resisting arrest, forcible rape, and disturbing the peace. The control group obtained consisted primarily of thieves and those convicted of nonviolent sex offenses, e.g., homosexual activities. The ages ranged from 19 to 62 with a mean of 41.

Group IV, consisting of normals, was included as a second control to provide a group so unlike the assaultive (and nonassaultive) criminal groups that it could be ascertained if the scales had the power to measure gross differences in control and hostility. This group of normals consisted of men tested as part of the follow-up in the Oakland Adolescent Growth Study, a longitudinal project being carried out by the Institute of Human Development, Berkeley, California. Entire lower-middle class high school classes were selected in the 1930s (in the same county in which the criminals reside) and have been studied up to the present (Jones, 1933). The MMPIs were obtained within the last 2 or 3 years when the subjects ranged in age from 36 to 39 with a mean of 37.

All random selection was done by means of a table of random numbers. No subject whose MMPI was incomplete was included, and all the MMPI profiles selected were later judged to be valid in that no indications of random responding were present.

It was felt that these groups should be ideal for the cross-validation of any scales designed to measure the expression and control of hostility. The members of two of the groups (I and II), had clearly demonstrated in their behavior both sufficient hostility and insufficient control to lead them to attack other people with degrees of violence ranging from battery to murder. Group III is a group of criminals who have never been charged with such behavior. Group IV is a group of normal subjects who are not known to have been convicted of any serious crimes and can therefore be presumed to be less hostile and more controlled as a group than any of the criminal samples. Thus it is to be expected that if the experimental scales measure

TABLE 3

VALUES OF *t* AND SIGNIFICANCE LEVELS FOR ALL COMPARISONS ON ALL SCALES

Scale	I versus II	I versus III	I versus IV	II versus III	II versus IV	III versus IV	I and II versus III	I and II versus IV	I, II, and III versus IV	I and II versus III and IV
L	1.026	0.026	2.604*	1.286	4.649**	2.986**	0.897	4.111**	3.863**	2.864**
F	1.372	0.459	2.341*	1.921	0.778	2.680*	1.623	1.510	2.098*	0.334
K	1.300	0.057	0.539	1.553	2.437*	0.480	1.130	1.853	1.358	1.803
Hs	1.586	1.344	1.552	0.363	0.438	0.031	0.359	0.360	0.225	0.450
Hs + .5*K*	0.987	1.411	1.674	0.502	0.735	0.208	1.012	1.269	0.932	1.399
D	0.688	1.026	2.854**	1.815	1.866	3.913**	1.841	2.513*	3.322**	0.209
Hy	0.740	0.955	1.122	1.948	2.117*	0.274	1.818	2.041*	1.495	2.354*
Pd	1.165	0.237	4.324**	1.102	3.250**	4.668**	0.716	4.314**	5.104**	1.999
Pd + .4*K*	0.572	0.216	4.212**	0.381	4.440**	4.659**	0.176	5.377**	5.967**	2.774**
Mf	0.263	0.878	0.394	1.269	0.091	1.500	1.337	0.235	0.835	0.697
Pa	1.236	0.541	2.289*	0.687	1.219	1.771	0.202	1.859	1.975	0.893
Pt	1.267	0.646	1.306	0.872	0.369	0.678	0.317	0.289	0.466	0.033
Pt + 1*K*	0.693	0.989	1.995	0.313	1.526	1.265	0.677	1.996	1.920	1.623
Sc	1.023	0.599	3.274**	0.480	1.693	2.297*	0.062	2.327*	2.379*	1.224
Sc + 1*K*	0.365	0.831	4.309**	0.538	3.830**	3.094**	0.782	4.389**	4.072**	2.846**
Ma	0.516	0.313	0.567	0.265	1.295	1.070	0.039	1.186	1.307	0.749
Ma + .2*K*	0.098	0.245	0.476	0.181	0.703	0.861	0.244	0.737	0.918	0.310
Si	1.220	0.039	1.307	1.284	0.364	1.467	0.952	0.917	1.340	0.022
AHI	0.358	0.490	2.024	0.187	2.145*	1.903	0.371	2.593*	2.710**	1.742
FTI	0.184	0.236	0.167	0.090	0.404	0.454	0.182	0.374	0.472	0.113
Ap	2.354*	0.720	2.899**	1.300	0.990	2.007	0.552	2.062*	2.251*	0.749
Hc	0.707	0.397	2.045	1.307	1.389	2.752*	1.115	1.793	2.305*	0.235
Ho	1.829	0.394	1.795	1.580	0.508	1.340	0.913	0.425	0.857	0.361
Hv	1.007	0.615	0.813	0.439	0.313	0.155	0.029	0.142	0.165	0.065
Hy − 5	2.125	0.540	0.244	3.364**	2.598*	1.044	2.403*	1.392	0.359	2.280*
Im	0.845	0.029	0.545	1.162	2.066*	0.672	0.782	1.513	1.257	1.428
Jh	1.813	0.610	0.912	1.353	1.498	0.145	0.642	0.580	0.303	0.756
Eo	0.720	0.332	2.058*	0.590	3.675**	3.363**	0.219	3.462**	3.917**	2.222*
Nu	1.227	0.028	0.305	1.531	1.513	0.337	1.008	0.807	0.383	1.148
Bc	0.322	0.243	0.782	0.668	0.362	1.277	0.583	0.570	0.900	0.031

Note.—The significance levels reported take into account heterogeneity of variance when necessary.
* $p \geq .05$.
** $p \geq .01$.

what they claim to measure, some significant differences between these groups should be obtained.

RESULTS

The data for both the clinical and experimental scales were subjected to a series of *t* tests. Both assaultive groups were compared with the nonviolent criminal group (I versus III and II versus III) and then both assaultive groups combined were compared with the nonviolent criminal control group (I and II versus III). This process was repeated using the normal control group (I versus IV, II versus IV, and I and II versus IV). Then the combined criminal groups were compared to the normal sample (I, II, and III versus IV). Finally all assaultive subjects were combined and tested against all nonviolent subjects (Groups I and II versus III and IV). For the sake of completeness the two criminal groups were compared (I versus II) as well as the two nonviolent groups (III versus IV).

The means and standard deviations for all groups on the experimental as well as the clinical scales are presented in Table 2 and the results of the *t* tests of the above comparisons are presented in Table 3.

DISCUSSION

Representativeness of the Samples

An examination of the results for the standard MMPI scales as presented in Tables 2 and 3 indicates that the criminal samples are representative of the general population of tested criminals. Their profiles are quite close together and are clearly differentiated from that of the normal sample. In addition the patterns of the criminal groups are similar to those reported in the literature. Since many of these men were in jail, Dahlstrom and Welsh's (1960) report of Beall and Panton's (1957) findings is most pertinent. These authors found a mean Welsh code for male prisoners of 4′ 28 91763-50/. In the present samples Group I has mean code of 4278236-950/; Group II, 4358-27916 0/; and Group III, 42-86735910/. The coefficient of concordance for these data is .716 ($p = .001$). Thus the four criminal groups are very similar in their overall patterns. Furthermore it should be noted that the code for the normal group (IV) is 59 341726/ 08, none of the mean scale values exceeding ± 1 *SD* from a *T* score of 50. Thus the data indicate that the criminal and normal samples are representative of the general populations from which they were drawn.[2]

[2] In passing, it is important to note the finding attributed to Levy, Southcombe, Cramer, and Freeman (Dahlstrom & Welsh, 1960, p. 308) that violent criminals differ from the nonviolent in the amount of elevation of *Pd* is certainly not borne out. (A reading of the Levy et al., 1952, article shows there is no statement to this effect, nor do their data seem to support such a conclusion, since both violent and nonviolent criminals appear to have equal *Pd* scores.)

Clinical Scales

An examination of Table 3 shows that while several of the standard scales discriminate between criminals and noncriminals, none is able to discriminate among the criminal groups (I and/or II versus III).[3] Nevertheless the *Hy* (Hysteria) scale is promising in this regard. The two violent groups are significantly higher than the two control groups on this scale. In addition, the comparison of the assaultive criminals with the nonviolent group (I and II versus III) approaches significance ($p < .10$). This is in agreement with Welsh and Sullivan's 1952 findings (Dahlstrom & Welsh, 1960, p. 190) that people with 43 (*Pd-Hy*) profiles while generally inhibited, occasionally lash out and express their chronic hostile feelings directly and intensely.

Experimental Indices

Examination of the experimental indices, Active Hostility Index (AHI) and Frustration Tolerance Index (FTI), shows that the former is capable of significantly discriminating criminals from noncriminals but is unable to differentiate assaultive from nonassaultive criminals. Since AHI is the sum of the *T* scores on *Pd* and *Ma* and the *Pd* scale is a highly significant discriminator, this is not a surprising result. While this index may be a good screening device in the general population (although poorer than *Pd* alone), it does not appear useful in differentiating between an assaultive criminal and his peaceable cellmate.

The FTI makes no significant discriminations. This is partly due to the pattern of scores on *Hy* noted above. It is possible that this index may be a predictor rather than a postdictor, since *D* may rise after apprehension and conviction.

Experimental Scales

The results on the experimental scales are, on the whole, discouraging. Some of the scales do not discriminate anything, others separate crimi-

[3] It is worth noting that the addition of *K* to *Pd* and *Sc* improves the discriminations between criminals and normals despite the fact that it was derived from a much different population.

nals from noncriminals, but only one is able to separate the assaultive from the nonviolent criminals. No significant discriminations were found on Ho, Hv, Jh, Nu, and Bc. In the case of Ho (Hostility), Hv (Overt Hostility), and Jh (Judged Manifest Hostility) this result certainly calls into question the validity of these scales. For the others it is a moot point.

The Nu (Neurotic Undercontrol) scale is meant to identify only subjects who are both neurotic and undercontrolled (see Table 1). While the criminal subjects appear undercontrolled, it cannot definitely be stated that they are neurotic as well. Therefore the present study cannot be considered a crucial test of this scale. The same reasoning applies to the Bc (Bimodal) scale which is supposed to identify subjects who are sometimes overcontrolled and at other times undercontrolled. While it can be stated that these subjects were undercontrolled at least once, it is not certain that they are often overcontrolled as well. However, since the assaultive groups were higher on *Hy*, Hy-5, Hc, and Eo we suspect that they may in fact be generally overcontrolled and lash out only on occasions. If this is the case, we would expect the Bc scale to make better discriminations.

In the case of Im (Impulsivity) the data indicate that the scale as described by Dahlstrom and Welsh (1960) is of questionable validity. However it should be noted that this scale is a shortened version of Gough's CPI scale (see Table 1). Since there is one barely significant result and a slight trend for the two nonviolent groups to be higher on the impulsivity scale, it is possible that the longer scale could better discriminate among the groups. There is a reversal [4] in the results, though, the two nonviolent groups scoring higher than the assaultive criminals.

Three other scales, Ap (Adjustment to Prison), Hc (Hostility Control), and Eo (Ego Overcontrol), separate the criminals from the noncriminals but do not separate assaultive criminals from nonassaultive criminals. Thus it is not clear whether these scales are measuring hostility control or whether they are detecting only general maladjustment or criminality. Ap produces significant differences on several comparisons and has the distinction of being the only scale on which normals score in the direction of less hostility or more control than the criminals. Two comparisons on Hc reach significance but each time in the reverse direction. Likewise, in the case of Eo, where the clearest discriminations are obtained, there are reversals, the scale attributing more control to the assaultive groups. It should be noted that the failure of the Eo scale to differentiate assaultive from nonassaultive criminals does not necessarily suggest invalidity, since there is no reason to suppose that one group of criminals has more control in a general sense than another group of criminals. However, it is surprising that the criminals should all be measured as having more control than the normals.

Finally, there is one scale, Hy-5 (Inhibition of Aggression), which apparently differentiates the combined assaultive criminal groups from the nonassaultive criminal group (I and II versus III) and from the combined control groups (I and II versus III and IV). Once again there is a reversal since the mean of the combined assaultive groups (I and II) exceeds that of the nonassaultive groups (III and IV). Moreover if the means of the groups are examined, we find that their order is Battery > Normals > Extreme Assaultive > Nonviolent criminals. Thus one would conclude that Battery cases inhibit their aggression more than normals, and murderers more than thieves. From a practical standpoint this would have unfortunate results, for not only is individual prediction impossible, but also in a probation setting it could mean releasing the men most likely to attack someone while retaining in prison those who are relatively harmless. This possibility points up the need for careful research and cross-validation before these or any scales are used in clinical or probation practice.

Despite the generally negative findings, there is a consistent trend: on those scales where significant results were obtained, the nonviolent groups appear less controlled or more hostile than the aggressive groups. One possible explanation for the reversals is that the criminal subjects, applicants for probation, are "faking good" in order to avoid prison. This hypothesis receives some support from the facts that the criminals are significantly higher than the normals on the *L* scale and that their average *F-K* indices are in the minus or "fake good" direction. However, the criminals were sufficiently frank in answering to obtain elevated *Pd* scores as well as elevations on many other scales. In addition, the *F-K* indices for the normal control group are just as strongly in the minus direction as those for

[4] The word *reversal* as used in this paper means a pattern of scores which indicates greater impulsivity or hostility or less control for the two nonviolent groups than for the two groups of assaultive criminals.

the criminals. Consequently neither the insignificant findings nor the obtained direction of the significant findings can be attributed solely to faking on the part of the criminals. Further, even if the results, both significant and insignificant, can be attributed to dissimulation, it seems unlikely that such easily faked scales have great utility.

If dissimulation did not produce these reversals on Hy-5, Im, Eo, Hc, and *Hy*, how else may these data be explained? An alternative hypothesis is that the extremely assaultive person is often a fairly mild-mannered, long suffering individual who buries his resentment under rigid but brittle controls. Under certain circumstances he may lash out and release all his aggression in one, often disastrous, act. Afterwards he reverts to his usual overcontrolled defenses. Thus he may be more of a menace than the verbally aggressive, "chip-on-the-shoulder," type who releases his aggression in small doses. If this is the case it means that the detection of assaultiveness is a difficult diagnostic task since we must separate the overly controlled but potentially aggressive person from the well-controlled nonaggressive person. The possibility that assaultive criminals may at different times be both overcontrolled and undercontrolled presents a particular problem to those who would construct scales on a rational rather than an empirical basis.

The existence of the trend just discussed should not obscure the basic finding that the experimental scales of hostility and control do not produce the expected differences between groups. In the cases of Ap, Im, Hc, Hy-5, and Eo there are some significant although often inconsistent results, but for the other five scales, not one of the 50 comparisons reached the .05 level of confidence. Why is this the case? Perhaps they measure something other than what their names imply. Or the cross-validation may have been insufficient. (In at least two cases apparently no cross-validation was attempted.) Another possibility is that they are so restricted in scope that the general title given to them is a misnomer. For example, the Hv (Overt Hostility) scale may be a fine measure of overt hostility as interpreted by a therapist in a permissive atmosphere, but it does not seem to have much relation to the overt hostility involved in the act of one of our subjects who bit off a man's nose in a fight. Thus Overt Hostility appears to be too general a description of the scale. A final possibility that cannot be discarded is that these five scales measure nothing at all.

Certainly the data indicate it is difficult if not impossible to identify an assaultive individual with reasonable accuracy using the MMPI scales now available. However there is some hope that the MMPI may have the potential to detect assaultive people since some items seem to be discriminating.

SUMMARY

Ten scales and two scoring indices of the MMPI which purport to measure hostility and control were cross-validated using four groups known on the basis of past behavior to differ on these dimensions. Group I, the most hostile and uncontrolled group consisted of 14 men convicted of Assault with a deadly weapon, Mayhem, Voluntary manslaughter, or Murder. Group II, moderately hostile and uncontrolled, consisted of 25 men convicted of Battery. Group III was a control group of 25 nonviolent criminals and Group IV a control group of 25 normals. Despite the wide variation in subjects, five scales and one index, Ho, Hostility; Hv, Overt Hostility; Jh, Judged Manifest Hostility; Nu, Neurotic Undercontrol; Bc, Bimodal Control; and FTI, Frustration Tolerance Index, showed no significant differences whatsoever in 60 comparisons. Five other scales and one index, Ap, Adjustment to Prison; Hc, Hostility Control; Hy-5, Inhibition of Aggression; Im, Impulsivity; Eo, Ego Overcontrol; and AHI, Active Hostility Index, showed some ability to separate criminals from noncriminals, but were generally unable to discriminate between the assaultive and nonassaultive criminal groups. An unexpected finding was that the great majority of the significant differences on the scales were reversals, the data indicating that assaultive criminals were better controlled and less hostile than the nonviolent criminals or the normals.

REFERENCES

BEALL, H. S., & PANTON, J. H. Development of a prison adjustment scale (PAS) for the MMPI. Unpublished manuscript, Central Prison, Raleigh, North Carolina, 1957. Cited by W. G. Dahlstrom & G. S. Welsh (Eds.), *An MMPI handbook: A guide to use in clinical research and practice*. Minneapolis: Univer. Minnesota Press, 1960. P. 308.

BLOCK, J. The development of an MMPI based scale to measure ego control. University of California, Berkeley, Department of Psychology, 1955. (Ditto)

COOK, W. W., & MEDLEY, D. M. Proposed hostility and pharasaic virtue scales for the MMPI. *J. appl. Psychol.*, 1954, **38**, 414–418.

DAHLSTROM, W. G., & WELSH, G. S. *An MMPI handbook: A guide to use in clinical research and practice*. Minneapolis: Univer. Minnesota Press, 1960.

GOUGH, H. G. *California Psychological Inventory manual*. Palo Alto: Consulting Psychologists Press, 1960.

HARRIS, R. H., & LINGOES, J. C. Subscales for the MMPI: An aid to profile interpretation. San Francisco: University of California School of Medicine, 1955. (Mimeo)

JONES, H. E. Personality study of adolescence. In, *Fourth conference on research on child development*. Washington: National Research Council, Division of Anthropology and Psychology, 1933. Appendix I.

LEVY, S., SOUTHCOMBE, R. H., CRAMER, J. R., & FREEMAN, R. A. The outstanding personality factors among the population of a state penitentiary: A preliminary report. *J. clin. exp. Psychopathol.*, 1952, **13**, 117–130.

MCGEE, SHANNA. Measurement of hostility: A pilot study. *J. clin. Psychol.*, 1954, **10**, 280–282.

PANTON, J. H. Predicting prison adjustment with the MMPI. *J. clin. Psychol.*, 1958, **14**, 308–312.

SCHULTZ, S. D. A differentiation of several forms of hostility by scales empirically constructed from significant items on the MMPI. *Pa. State U. Abstr. Doct. Dissertations*, 1954, **17**, 717–720.

SIEGEL, S. M. The relationship of hostility to authoritarianism. *J. abnorm. soc. Psychol.*, 1956, **52**, 368–372.

STATE OF CALIFORNIA. *Penal code of the State of California*. San Francisco: Bancroft-Whitney, Bender-Moss, 1959.

(Received August 2, 1961)

Journal of Abnormal and Social Psychology
1962, Vol. 65, No. 6, 438–442

ACCEPTANCE BY OTHERS AND ITS RELATIONSHIP TO ACCEPTANCE OF SELF AND OTHERS:

A REPEAT OF FEY'S STUDY

JOHN E. WILLIAMS [1]

Wake Forest College

One of the more frequently studied problems in the area of self-concept research has been the relation of acceptance of self to acceptance of others. In the studies of this problem reviewed by Wylie (1961, pp. 235–238), it is reported with regularity that these two phenomenological variables are positively associated, i.e., persons who are high in self-acceptance tend to be accepting of other persons. This finding is consistent with the expectations of several personality theorists (e.g., Rogers, 1951) who have proposed that an individual's attitudes towards others are determined, to a considerable degree, by his attitudes toward himself.

Less research attention has been directed to the relation of the variables of acceptance of self, and acceptance of others, to the nonphenomenological variable of actual acceptance by others. McIntyre (1952) studied this problem and was unable to find any relationship between either acceptance of self, or acceptance of others, and actual sociometric acceptance in male college students. These results were interpreted as not being in support of Rogers' (1951) expectations that ". . . the person who accepts himself will, because of this self-acceptance, have better interpersonal relations with others" (p. 520).

Fey (1955) was prompted by McIntyre's (1952) negative findings to explore the notion that the critical determinant of good interpersonal relations is not the absolute level of self-acceptance or acceptance of others but, rather, the interrelationship of the two variables. Fey tested this hypothesis with a specially constructed questionnaire administered to a group of third-year medical students. His hypothesis was confirmed when the difference score computed between the acceptance of self score and the acceptance of others score was found to have a small but significant negative correlation ($r = -.27$) with the sociometric score reflecting acceptance by others. Fey interpreted this result as indicating that persons who are relatively more accepting of self than of others tend to be less well liked by others. Fey also found, as had McIntyre, that when acceptance of self, and acceptance of others, were considered separately, they were not related to acceptance by others.

[1] The author wishes to thank Barbara Hathcock Turner for her assistance in the data analysis. Gratitude is also expressed to the officers and members of the local chapters of the Kappa Alpha and Lambda Chi Alpha fraternities for their participation in this investigation.

In addition to providing scores relating to self-acceptance, and acceptance of others, the questionnaire used in the Fey (1955) study also provided a score purporting to deal with the individual's perception of his acceptance by others. This score, called Expected Acceptance (EA), was found to have significant positive correlations with the scores indicating acceptance of self and acceptance of others. On the other hand, when the EA scores were compared with the actual sociometric acceptance scores, a zero correlation ($r = .00$) was obtained. This result, taken at face value, would seem to indicate that, for Fey's medical school students, there was *no* relation between perceptions of acceptance by others and actual acceptance by others. This remarkable finding, which Fey passed over rather lightly, seemed to warrant further study. Certainly it was felt, there must be at least a moderate, positive relationship between a person's perception of acceptance and his actual acceptance by others.

One possible explanation of Fey's (1955) puzzling result was methodological in nature. Fey's instructions for determining actual acceptance by others involved asking each subject to pick the five of his fellow students which he "liked best—just as persons." Each subject's Actual Acceptance score represented the number of times he was chosen in this way by his peers. In comparison with this definition of Actual Acceptance (AA), the five items making up the EA scale of the Fey questionnaire were, as follows: Number 4, "people are quite critical of me"; Number 15, "I feel 'left out,' as if people don't want me around"; Number 19, "people seem to respect my opinion about things"; Number 37, "people seem to like me"; Number 45, "most people seem to understand how I feel about things." From the point of view of content, item Number 37 would seem to be rather directly related to "being liked—just as a person" and item Number 15 might be taken as an inverse indicator of the same variable. On the other hand, the content of items Number 4, Number 19, and Number 45 would not seem, necessarily, to be related to "being liked" in any straightforward manner. On the other hand, Fey reported that the five-item EA scale had a split-half reliability of .89 which indicates that the items do tend to vary together. Thus, while the EA scale appears to measure something, it remains a somewhat equivocal measure of "expected acceptance by others."

It was thought that a simple and perhaps illuminating addition to Fey's (1955) procedure would be to ask subjects to predict the number of their fellow subjects who would choose them as being among the five persons they liked best. Such a score might be called Predicted Acceptance (PA) to distinguish it from Fey's EA score. It was felt that this new procedure might provide some clarification of Fey's finding of zero correlation between "expected" and "actual" acceptance.

METHOD

Subjects

Subjects were 74 members of two social fraternities at Wake Forest College. Subjects from two fraternities were used in order to have an adequate number of subjects while still dealing with groups sufficiently small that sociometric procedures remained meaningful. The great majority of subjects lived in the dormitory section assigned to their fraternity and which housed the fraternity lounge and meeting rooms.

To insure that all persons in each fraternity group had had an opportunity to get to know one another well, only those men were included who had been full members of the fraternity for at least one year prior to the time of the study. Thirty-six of the subjects were from Fraternity A representing 100% of the members meeting the above restriction; 38 of the subjects were from Fraternity B representing 95% of the eligible members.

Testing Procedure and Questionnaire

The questionnaires were administered by the author, a college faculty member, to each fraternity group separately at their regular evening meeting times. Members who were not present at the meeting were contacted shortly afterward and were administered the questionnaire individually.

The questionnaire employed was the same 45-item instrument developed by Fey (1955) to measure Acceptance of Self (AS), Acceptance of Others (AO), and expected Acceptance by Others (EA). The questionnaire presented the subject with statements such as: "I'm pretty satisfied with the way I am" (AS), "I like people I get to know" (AO), and "People are quite critical of me" (EA). There were 20 AS items, 20 AO items, and 5 EA items. The subject rated these items on a five-point scale from very rarely to almost always.

After completing Fey's (1955) questionnaire, the subject was presented with a numbered, alphabetical list of the eligible members of his fraternity. The subject was asked to indicate, by means of the number beside each name, the "five of your fraternity brothers whom you like best—just as persons." Following this, an additional procedure, not used by Fey, was employed. The subject was requested to consider again the list of his fraternity brothers and to answer the question, "How many do you think will list you among the five they

like best?" The subject did not sign his name to his paper but used his number in its place.

RESULTS

The first step in the analysis of the results was to determine whether there were any differences between the two fratrnity groups on the measures under consideration. Means and variances were examined for each of the measures used. No sizable differences were found and, thus, the data for the two groups were pooled for all subsequent analyses.

Mean Differences between Fey's Group and the Current Group

Table 1 provides a summary of the means and standard deviations for the primary measures from Fey's (1955) study and from the current study. An analysis by t tests indicated that certain of the observed mean differences were statistically significant. Subjects in the present study had a lower AS score ($p < .001$) and a lower AO score ($p < .001$), while Fey's subjects had a lower AA score ($p < .001$). There was no significant difference in the EA score. When certain discrepancy scores used by Fey were examined, it was found that subjects in the current study had a significantly greater AS-AO mean discrepancy score ($p < .01$) and a significantly smaller ($p < .01$) mean discrepancy score between EA and AA.

Correlations among Scales

Product-moment correlation coefficients were computed for each pair of the following scores: AS, AO, EA, PA, AA, and the AS-AO discrepancy score. These coefficients are reported in Table 2 together with the corresponding coefficients obtained by Fey (1955).

The analysis summarized in Table 2 revealed that the three questionnaire scales—AS, AO, and EA—were all substantially interrelated in a positive manner, as Fey (1955) had reported. The AS-AO difference score was positively related to the AS score, as Fey had found, but was not significantly related to the AO score ($r = -.12$), where Fey had reported a significant negative correlation ($r = -.55$). Consistent with Fey, a small but significant positive relationship was found between AS-AO and EA scores.

From the Table 2 correlations involving the AA score, it can be seen that, consistent with Fey's (1955) findings, this score was not significantly related to AS or AO. However, no evidence was found supporting Fey's finding of a negative association between AA and AS-AO discrepancy score. An examination of the correlations involving AA and the two "perception of acceptance" scores revealed that the EA score from Fey's questionnaire showed a weak positive association with AA ($r = +.25$) while the new PA score showed a substantial positive correlation with AA ($r = +.68$). On the other hand, the PA score was not related significantly to any of the questionnaire scores—AS, AO, EA, or AS-AO.

TABLE 1

SUMMARY OF PRIMARY MEASURES FROM FEY'S STUDY AND THE PRESENT STUDY

Scale	*M*		Ranges		*SD*	
	Fey	Williams	Fey	Williams	Fey	Williams
Acceptance of Self (AS)	80.62	71.08	55–95	45–92	7.90	11.39
Acceptance of Others (AO)	75.53	61.55	57–96	44–82	8.39	8.39
Expected Acceptance by Others (EA)	19.53	18.55	14–24	11–24	4.19	3.07
Predicted Acceptance by Others (PA)	—	7.59	—	0–30	—	5.61
Actual Acceptance by Others (AA)	4.19	4.94	0–10	0–17	2.32	3.64

TABLE 2

CORRELATIONS AMONG SCALES

Scale	AO	AS-AO	EA	PA	AA
AS	.64** (.43)**	.68** (.49)**	.62** (.71)**	.11	.14 (−.07)
AO		−.12 (−.55)**	.48** (.43)**	.14	.19 (.20)
AS-AO			.34** (.26)*	.02	−.01 (−.27)*
EA				.14	.25* (.00)
PA					.68**

Note.—Fey's coefficients are in parentheses.
* $p < .05$.
** $p < .01$.

DISCUSSION

Although this study was intended to serve as a replication of Fey's (1955) study, the data in Table 1 indicate that, in a strict sense, it cannot be so considered. The mean score differences between Fey's medical school students and the undergraduate fraternity men in the current study strongly suggest that two populations, with rather different group characteristics, were involved in the two studies. Relative to the fraternity men in the current study, Fey's medical school students had higher AS, AO, and

a smaller AS-AO discrepancy. In addition, the medical school students showed a lower than mean AA score.

This last result was due, apparently, to the failure of a number of Fey's (1955) subjects to follow instructions and name *five* persons in the sociometric procedure from which the AA scores were obtained. If all subjects had named five persons, the mean AA score would have been 5.00 while Fey reported a mean AA score of only 4.19. This result might be interpreted as indicating that a considerable number of Fey's subjects were unable to name as many as 5 of their 57 fellow medical students whom they "liked best—just as persons." If this was true, it would suggest that these 58 third-year medical students were not sufficiently well acquainted with one another "just as persons" to make the procedure for determining AA scores very meaningful. An additional result of Fey's study, which might be interpreted in the same light, was his failure to find any relationship between his EA and AA scores. If Fey's subjects did not know one another sufficiently well to make meaningful sociometric choices, it would not be surprising if their AA scores were not related to their expectations of acceptance by others.

One does not have to agree with the above speculations to grant that groups of undergraduate fraternity men, all active members for at least one year, do have a good basis for judging personal likes and dislikes among their fraternity brothers. Almost all of these fraternity men were able to name at least five of their brothers whom they "liked best—just as persons" as was seen in the mean AA score of 4.94 as compared to the theoretical value of 5.00. It is also interesting that, in the current study, the EA scores were found to be related ($r = .25$, $p < .05$) to the AA scores. This indicates that, even with this somewhat questionable measure of expected acceptance, there was a significant positive relationship between subjects' expectations of acceptance and actual acceptance by their peers.

The foregoing argument implies that the current study can be considered a good—if not superior—testing ground for Fey's (1955) hypothesis. With this in mind, it can be noted that no evidence was found in support of Fey's hypothesis that the AS-AO discrepancy score would be negatively related to the AA score. Thus, while the results of the current study were in agreement with Fey's conclusion that "persons who think relatively much better of themselves than they do of others (high AS-AO) tend to feel accepted by others (EA)," there was no evidence that these same persons "actually are significantly less well liked by (others)" (p. 276).

The results of the study provided some clarification regarding the relationship of perceived acceptance to actual acceptance. As noted, the EA score showed a low, positive correlation ($r = .25$, $p < .05$) with the AA scores. On the other hand, the new PA score was shown to have a sizable positive correlation ($r = .68$, $p < .001$) with the AA scores. Taken together, these correlations indicate that, in the current study, there was general agreement in the subject's perceptions of acceptance and his actual acceptance by his peers. Taken separately, the correlations suggest that that PA procedure is superior as a predictor of actual acceptance by others.

The PA and EA scores appeared to have differing characteristics apart from their relationship to the AA scores. It was seen in Table 2 that EA correlated significantly with the AS and AO scores while PA did not. EA appears to be another phenomenological variable within a general subjective frame of reference. PA, on the other hand, seems not to share this same frame of reference and its intercorrelations with other scores are much more similar to the AA intercorrelations than to those involving the questionnaire scores. Fey (1955) used the phrase "feel accepted" to describe the meaning of high EA scores. Following this terminology, the current study suggests that one's "feelings of acceptance" and one's "predictions of actual acceptance" are quite different with the former related more closely to other subjective feelings about self and others, and the latter more closely related to interpersonal reality.

In general, the results of this study seem to return the matter to McIntyre's (1952) original conclusions. While it was demonstrated in this study that subjects can predict their acceptance by others, no evidence was found that acceptance by others is related to acceptance of self, acceptance of others, or the interrelationship of these two variables.

SUMMARY

This study repeated an investigation of Fey (1955) who gave a questionnaire and sociometric procedure to third-year medical students and reported that AS and AO were not related to AA but that the AS-AO discrepancy score was negatively correlated with AA. Fey also reported that the EA score from his question-

naire was not related to actual acceptance. Seventy-four undergraduate fraternity men repeated Fey's procedure and, in addition, made a direct estimate (PA) of their expected acceptance by their peers. While the results of the study were similar to Fey's with regard to the positive interrelationships of the questionnaire scores, no evidence was found in support of his hypothesis that the AS-AO discrepancy score would be related to acceptance by others. It was found that Fey's EA score and the new PA score both were positively correlated with the AA scores, with the PA score having the higher correlation. It was suggested that the AA measure may not have been too meaningful for Fey's subjects. It was concluded that subjects were able to predict their actual acceptance by their peers, but there was no evidence that actual acceptance was related to the questionnaire measures of acceptance of self, acceptance of others, or the discrepancy between these two variables.

REFERENCES

FEY, W. F. Acceptance by others and its relation to acceptance of self and others: A revaluation. *J. abnorm. soc. Psychol.*, 1955, **50**, 274–276.

MCINTYRE, C. J. Acceptance by others, and its relation to acceptance of self and others. *J. abnorm. soc. Psychol.*, 1952, **47**, 624–626.

ROGERS, C. R. *Client-centered therapy*. Boston: Houghton Mifflin, 1951.

WYLIE, RUTH C. *The self concept*. Lincoln: Univer. Nebraska Press, 1961.

(Received August 23, 1961)

"The new in tests. . . . . "

How would you like to use a projective personality test which can be scored quickly and objectively? Whose interpretation is based on careful factor-analytic studies? Which has norms for 11-year-old through adult levels (over 1900 selected Ss)? Uses an interesting and disarming cartoon-like format?

A NEW TEST!!

Cassel and Kahn

GROUP PERSONALITY PROJECTIVE TEST

A NEW TEST!!

Order Examiner's Set including general manual, scoring keys, 10 test booklets, and 100 answer sheets, enough material for a good-sized testing program (and only $13.50).

•

For further information, write to

PSYCHOLOGICAL TEST SPECIALISTS
Box 1441
Missoula, Montana

"The new in tests. . . . . "

Chestnut Lodge Symposium

PAPERS PRESENTED ON THE 50th ANNIVERSARY, 1910–1960

Contents

DEXTER M. BULLARD: Introduction

JOHN C. WHITEHORN: Alienation and Leadership

DAVIC McK. RIOCH: The Sense and the Noise

ALFRED H. STANTON: Milieu Therapy and the Development of Insight

DON D. JACKSON and JOHN H. WEAKLAND: Conjoint Family Therapy

ROBERT A. COHEN and MABEL BLAKE COHEN: Research in Psychotherapy

CLARENCE G. SCHULZ: Case Report of an Obsessional Patient

OTTO ALLEN WILL, JR.: Paranoid Development and the Concept of Self: Psychotherapeutic Intervention

HAROLD F. SEARLES: Sexual Processes in Schizophrenia

DONALD L. BURNHAM: Identity Definition and Role Demand in the Hospital Careers of Schizophrenic Patients

$3.50

William Alanson White Psychiatric Foundation

1610 New Hampshire Ave., N.W., Washington 9, D.C.